Summer 8-22-2020

Spruce Budworm Defoliation Detection and Host Species Mapping Using Sentinel Satellite Imagery

Rajeev Bhattarai
University of Maine, rajeev.bhattarai@maine.edu

SPRUCE BUDWORM DEFOLIATION DETECTION AND HOST SPECIES MAPPING USING SENTINEL SATELLITE IMAGERY

By Rajeev Bhattarai

B.S. Tribhuvan University, Nepal, 2017

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Forest Resources)

The Graduate School

The University of Maine

August 2020

Advisory Committee:

Parinaz Rahimzadeh-Bajgiran, Assistant Professor of Remote Sensing of Natural Resources, The University of Maine, Orono, Maine, Advisor

Aaron Weiskittel, Professor of Forest Biometrics and Modeling, The University of Maine, Orono, Maine

David MacLean, Emeritus Professor of Forest Ecology, The University of New Brunswick, Fredericton, Canada

Brian Roth, Associate Director of SeedTree, Maine

SPRUCE BUDWORM DEFOLIATION DETECTION AND HOST SPECIES MAPPING USING SENTINEL SATELLITE IMAGERY

By Rajeev Bhattarai

Thesis Advisor: Dr Parinaz Rahimzadeh-Bajgiran

An Abstract of the Thesis Presented

in Partial Fulfillment of the Requirements for the

Degree of Master of Science

(in Forest Resources)

August 2020

Insects are one of the most significant agents causing landscape level disturbances in North American forests, and among them, spruce budworm (*Choristoneura fumiferana*; SBW) is the most destructive forest pest of northeastern Canada and U.S. The SBW occurrence, its damage extent and severity are highly dependent on characteristics of the forests and availability of the host species (spruce (*Picea* spp.) and balsam fir (*Abies balsamea* (L.) Mill.)). This study developed novel methodologies to detect and classify SBW defoliation and to map SBW host species using remote sensing techniques. Optical multispectral remote sensing satellite imagery presents a valuable data source for regional-scale mapping of forest composition as well as defoliation severity and can be effectively used for monitoring insect outbreaks. This study developed two separate models to map both the distribution and abundance of SBW host species as well as the severity of defoliation at 20 m spatial resolution utilizing Sentinel imagery. The two models were integrated to effectively monitor the SBW defoliation.

For the detection and severity classification of SBW defoliation, we used Sentinel-2 imagery and site variables (elevation, aspect, and slope) and compared the capabilities of various spectral vegetation indices (SVIs), in particular red-edge SVIs, to detect and classify SBW defoliation using Support Vector Machine (SVM) and Random Forest (RF) models. The study was carried out in the Northern part of New Brunswick, Canada. Results showed the superiority of RF in model building for defoliation detection and classification into three classes (non-defoliated, light and moderate) with overall errors of 17% and 32%, respectively. The most important Sentinel-2 based variables for the best model were Inverted Red Edge Chlorophyll Index (IRECI), Enhanced Vegetation Index 7 (EVI7), Normalized Difference Infrared Index 11 (NDII11), Modified Chlorophyll Absorption in Reflectance Index (MCARI), and Modified Simple Ratio (MSR). Elevation was the only site variable significant in the final model. The study concluded that red-edge SVIs were more effective variables for light defoliation detection compared to the traditional SVIs such as Normalized Difference Vegetation Index (NDVI) and EVI8. These findings can help improve the current remote-sensing based SBW defoliation detection techniques.

For SBW host species classification, Sentinel-1 synthetic aperture radar (SAR) and multi-spectral Sentinel-2 imagery were used in combination with several site variables (elevation, slope, aspect, topographic wetness index, soil types, projected climate site index for year 2030, and improved Biomass Growth Index (iBGI)). The study was carried out in the same location where the first study was conducted but extended to a larger area (northern parts of New Brunswick, Canada) using a total of 191 variables. We found Sentinel-2 time series in combination with single spectral bands and spectral vegetation indices (SVIs) promising to map SBW host species using a RF algorithm, with an overall accuracy (OA) of 71.34% and kappa

coefficient (*K*) of 0.64. The use of Sentinel-1 SAR data alone with elevation showed a decent result (OA: 57.5 and *K*: 0.47). Furthermore, the combination of Sentinel-1, Sentinel-2 and elevation provided us with an OA of 72.3% and *K* of 0.65. The most important Sentinel-2 variables for the best model were from the images of late spring and fall seasons including three single spectral bands and seven SVIs mostly from near-infrared, red-edge and shortwave-infrared regions. Prediction of spatially explicit SBW host species data is essential for identifying vulnerable forests, tracking the SBW defoliation and minimizing the forest loss as well as serving as a vital input for modelling and managing insect impacts at the landscape and regional scales.

DEDICATION

I dedicate this thesis to my parents (**Parameshwar Bhattarai** and **Sushila Wagle**) and my advisor **Dr. Parinaz Rahimzadeh-Bajgiran**.

ACKNOWLEDGEMENTS

First of all, I would like to express my heartfelt gratitude towards my advisor Dr. Parinaz Rahimzadeh-Bajgiran for believing in me and investing her time and resources on my study. Her motivations and the opportunities I was provided with to see a broader world of academia are invaluable for me. She accepted the crudest form of my knowledge and work, derived sense out of it and facilitated me to shape it to the ultimate form professionally. The door to Dr. Rahimzadeh's office was always open whenever I ran into trouble or had any questions about my research or writing. She consistently allowed me to work freely on my ideas but steered me in the right direction whenever she thought I needed it. Her kind-hearted nature instilled a comfort in me to perceive her not as my advisor but as a considerate and helpful friend, never ordering me to perform but sitting together and solving the problems no matter how minor they were. She was not only a responsible advisor but also a very good human to me. I always feel so lucky to have her as my advisor and confess that this work is a combined product of her and my hard work.

I extend my thankfulness to all of the committee members of my thesis, Dr. Aaron Weiskittel, Dr. David MacLean, and Dr. Brian Roth. It was a huge pleasure to have Dr. MacLean, one of the pioneer researchers of spruce budworm, helping me to evolve in this field. I appreciate Dr. Weiskittel for his statistical assistance during my research. He was always with me to monitor my progress throughout these two years and assist in improving when needed. Together with other members, I can't miss Dr. Roth, who deserves a very big "**thank you**" from me. He not only assisted me throughout the project but also provided an environment of home away from my home. He also helped me to increase my interactions with the forest stakeholders

of Maine. I am also very thankful to Dr. Rob Johns and Dr. Chris Hennigar for providing the field data for my work.

I also thank all the faculty and friends from the School of Forest Resources for being a wonderful family to me. I especially thank Sandesh Shrestha, Elias Ayrey, Catherine Chan, and Jeanette Allogio for being good friends during my early days in the USA. Besides, I thank Aaron Meneghini for helping me out with Sentinel-1 SAR image processing and interpretation.

Whatever I achieve in my life, I always remember the contribution of all the teachers, fathers and sisters of my high school, Notre Dame School, Bandipur, Nepal for inspiring me to pursue not only for academic excellence but also for being an admirable human. I would never have seen this entirely different world if Notre Dame wasn't there.

My family is imprinted on my heart and memory, and it is difficult to put a figure on the value they have in each and every achievements of mine. Words cannot fully express my gratitude to all the individuals who have contributed to my standing today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ARI = Anthocyanin Reflectance Index

ASMs = Aerial Sketch Maps

BF = Balsam Fir

BL = Broadleaved

ClRE = Chlorophyll Red Edge

CRI = Carotenoid Reflectance Index

EROS = Earth Resource Observation Satellite

ESA = Europian Space Agency

EVI = Enhanced Vegetation Index

GNDVI = Green NDVI

GRD = Ground Range Detected

iBGI = improved Biomass Growth Index

IRECI = Inverted Red Edge Chlorophyll Index

IW = Interferometric Wide

JERS = Japanese Earth Resource Satellite

K = Kappa Coefficient

MCARI = Modified Chlorophyll Absorption in Reflectance Index

MDA = Mean Decrease in Accuracy

MODIS = Moderate Resolution Imaging Spectroradiometer

MSR = Modified Simple Ratio

MTCI = MERRIS Terrestrial Chlorophyll Index

NB = New Brunswick

NDII = Normalized Difference Infrared Index

NDVI = Normalized Difference Vegetation Index

NIR = Near Infra-Red

OA = Overall Accuracy

OOB = Out of Bag

PA = Producer’s Accuracy

QC = Quebec

RF = Random Forest

RS = Remote Sensing

SAR = Synthetic Aperture Radar

SBW = Spruce Budworm

SNAP = Sentinel Application Platform

SPOT = Satellite Pour l’Observation de la Terre

SRTM = Shuttle Radar Topography Mission

SVIs = Spectral Vegetation Indices

SVM = Support Vector Machine

SWIR = Short Wave Infra-Red

S2REP = Sentinel-2 Red Edge Position

TOA = Top-Of-Atmosphere

UA = User’s Accuracy

VH = Vertical-Horizontal

VSURF = Variable Selection Using Random Forest

VV = Vertical-Vertical

WDRVI = Wide Dynamic Range Vegetation Index

CHAPTER 1

INTRODUCTION

The general purpose of this thesis is to provide a solution for the detection and severity classification of eastern spruce budworm (*Choristoneura fumiferana* Clem.; SBW) defoliation and mapping its host species in northeastern forests of America. Technically, this thesis aims to identify the significant remote sensing and site variables and algorithms to map SBW defoliation extent and severity as well as its host tree species. When SBW outbreak occurs, it has the potential to invade a wide extent of the landscape and therefore assessing SBW induced damage at ground level is difficult and uneconomical. At the same time, updated data on SBW tree host species are essential for the prediction of hot spots and better validation of SBW defoliation data generated from methods other than field surveys. In this thesis, the intent is to develop novel remote sensing (RS) based techniques capable of detecting and classifying SBW defoliation at the landscape scale, along with mapping its host availability at a resolution fine enough to be used for forest management. To this end, the newly-launched Sentinel-2 satellite data were used for SBW defoliation detection and severity classification. A combination of Sentinel-1 and Sentinel-2 satellite imagery was also utilized to map the SBW host species. Other site variables such as elevation, slope, aspect, Topographic Wetness Index (TWI), soil types, projected climate site index for year 2030, and improved Biomass Growth Index (iBGI) were also employed to assist with the modeling.

1.1 Background on Forest Pests and SBW Defoliation

Natural disturbances such as wildfire, pest outbreaks, windthrow and ice storms are constantly changing forest ecosystems around the world and in particular North America.

Among these, insects and pest outbreaks are the leading forest disturbance agents in North American forests impacting about 20.4 million hectares of area (average annual basis) in the United States and causing an average annual economic cost of $1.5 billion (Dale et al., 2001). The changing climate and human disturbances on the environment also exacerbate the situation for forest resources (Dale et al., 2001). The pests and parasites in the forest not only alter tree growth and timber production but also undermine wildlife habitat (Bennetts et al., 1996). The introduction of pests is capable of disturbing the spatial distribution of the tree species (Dale et al., 1991). Among various forest pests, SBW is one of the prominent native pests responsible for the widespread landscape-scale changes in structure and composition of boreal forests in northeastern America (MacLean, 1984; Morin, 1998; Rahimzadeh-Bajgiran et al., 2018).

The SBW is a native forest pest of eastern North America causing widespread tree mortality and loss of productivity in balsam fir (*Abies balsamea* (L.) Mill.) and spruce (*Picea* spp.) and the severity of damage during its periodic outbreaks is attributed to the availability of mature balsam fir (Blais, 1983).

The history of SBW outbreak in southern Quebec (QC) and New Brunswick (NB) in Canada over the past 250 years as studied by Blais (1983) and over the last 400 years as reviewed by Boulanger et al. (2012) shows the interval between outbreaks has decreased after the 20th century. As per the studies aforementioned, the intervals fell from a range of 42-75 years (before the 20th century) to 19-34 years during the 20th century. Looking at the history of SBW infestations, six SBW outbreaks occurred in this region before 1983, among which all were severe except the 1949 outbreak (Blais, 1983). The dendrochronological study of five old-growth stands and 12 heritage buildings from Boulanger et al. (2012) over the past 400 years showed the potential occurrences of nine outbreaks, including three additional uncertain ones. The eruptions

were restricted to a small area before the 20th century, but after that, the outbreaks were more widespread and severe (Kettela, 1983). A study in eastern Canada in 1980 showed that the young balsam fir and spruce species attained a similar degree of mortality as that of mature stands which was higher when compared with the previous outbreaks (Blias 1981). Following the past trend of cyclic SBW attack, the recent outbreak has already started in 2006 in Quebec defoliating over nine million ha of the forest by the summer of 2019 (Ministère des Forêts de la Faune et des Parcs. 2019). Also, the defoliation is moving southwards from Quebec towards the northern part of NB and the state of Maine in the U.S.A. (MacLean et al., 2019).

In addition to field surveys that are tedious and expensive to conduct, traditionally aerial surveys are used to detect and classify the severity of SBW defoliation at the landscape scale in North America. The overall accuracy of aerial surveys is variable based on the differences in mapping techniques in Canada and the U.S.A., which can range from a rather fine to coarse spatial resolutions (MacLean and MacKinnon, 1996; Hennigar et al., 2013). For the current outbreak starting in 2006 in the Quebec region, the aerial survey of defoliated forest area is continuously conducted on an annual basis by the Ministry of Forests, Fauna and Parks *-(Ministère des Forêts de la Faune et des Parcs) (Huang, 2015). Aerial surveys have already been started in NB also for the detection of the current outbreak (MacLean et al., 2019). Given that the existing SBW population has not yet caused defoliation in Maine, ground monitoring is currently conducted using field surveys only.

1.2 SBW Phenology and Defoliation Detection

SBW life stages per year include egg, larva (with six larval instars: young larvae (L1 and L2) and old larvae (L3 to L6)), pupa, and adult (Figure 1.1). In the case of eastern North America, SBW moths emerge from mid-July to early August, with no feeding of L1-L2 until

spring. From about early June to early July, larvae feed on the current-year shoots, if current-year shoots become depleted, the larvae will feed on older foliage (Royama, 1984). L5 and L6 instars of SBW consume about 95% of the host needles, while other larval instars consume less than 5% altogether (Miller, 1977). The SBW uses locations on various parts of trees like male flower cups, beneath buds and bark scales, amongst lichens or in bark crevices to hibernate in the form of tiny larvae, which emerge in spring as soon as the buds start to expand (Rose and Lindquist, 1985). The larvae then begin attacking the fresh needles, unopened buds or when available, male flowers. Later they feed on the expanding buds, and as the new shoots grow, they spin a fine silk-like structure between shoots. The feeding pattern goes from first new needles and then to old needles, and if heavy, the scorched reddish appearance of foliage can be easily detectable using aerial survey (Rose and Lindquist, 1985). During the feeding season of SBW, the branches and canopy of the defoliated trees gain a distinct red color. While feeding on balsam fir-spruce leaves, SBW forms a web-like feeding tunnel around several of the terminal twigs and partially consumed needles become entrapped in this webbing, then dry, and turn a reddish color (Leckie et al., 1992).

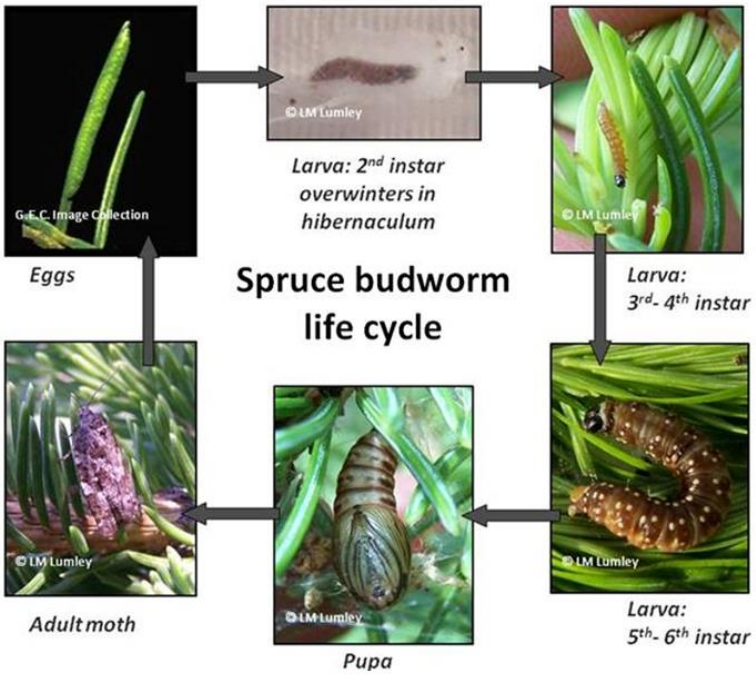


Figure 1.1. SBW life cycle (Credit: Michel Cusson, *http://rclevesque.ibis.ulaval.ca/en/home/research-area/begab-budworm-ecogenomics-applications-and-biotechnology/*).

SBW larvae strongly prefer to feed only on the current-year foliage of balsam fir and spruce, and feed on older age-classes of foliage only if all current-year foliage is consumed (Royama 1984). Generally, we can see two categories of defoliation related to this pest, current-year/annual defoliation and cumulative defoliation. Current-year defoliation is the amount of leaf damage occurring in a particular year, whereas cumulative defoliation is the result of multiple years of defoliation happening in an area. Current-year defoliation is estimated by determining the percentage of the current foliage age class that is consumed by SBW and is correlated with reddish discoloration of foliage. This provides information on the location and severity of defoliation and distribution of budworm populations in any year (Hall et al., 2016; Rahimzadeh-Bajgiran et al., 2018).

The optimum visual window to detect and map current-year defoliation is short and persists only for about 2-3 weeks. This time period, as mentioned above, relates to the redness of the defoliated needles, which often fall because of rain and wind or become less pronounced in color after about two weeks (MacLean and MacKinnon, 1996). Therefore any survey should be conducted during this time frame before the needles dry and fall off. Generally, the estimation of current-year defoliation on a landscape level is performed using aerial sketch maps (ASMs). The accuracy of ASMs depend upon various factors like observer's ability to recognize the defoliation severity and identify the location, timing of the surveys, weather conditions and nature of the defoliation (Kettela, 1982). The accuracy of the ASMs also can differ widely from person to person (Waters et al., 1958), and showed lower accuracy while detecting lower level defoliation than that of higher severity level (Taylor and MacLean, 2008). This technique also demands considerable human resource and is time-consuming (Huang, 2015).

1.3 Satellite Remote Sensing Sensors for Vegetation Studies

There is a wealth of both open source as well as commercial satellite remote sensing sensors such as passive optical multispectral/hyperspectral and active microwave sensors available to study vegetation characteristics and change detection. Landsat satellite series have a very long history of image availability from 1972 and have been used extensively for vegetation studies. These satellites fall into medium resolution optical multispectral systems making them very useful for vegetation studies at local and regional scales. The latest Landsat mission (Landsat 8) carrying Operational Land Imager and Thermal InfraRed Sensor provides data in 11 regions of the electromagnetic spectrum (11 bands) including bands in visible, near-infrared, short wave infrared and thermal regions providing data with a spatial resolution ranging from 15-100 m and a temporal resolution of 16 days. Moderate Resolution Imaging Spectroradiometer

(MODIS) is another passive optical multispectral satellite providing high temporal resolution (up to one to two days) but coarser spatial resolution (250 m-1000 m) data over 36 spectral bands (0.4µm-14.4µm). Among two MODIS satellites currently in space, MODIS-Terra was launched in 1999 while MODIS-Aqua was launched in 2002. Both Landsat and MODIS satellite imagery are free-of-charge and publicly available. Compared to Landsat, MODIS data are more suitable for landscape scale regional and global studies and have been used successfully for vegetation amount mapping and stress detection (Jones and Vaughan, 2010; Rahimzadeh-Bajgiran et al. 2012b).

Satellite Pour l'Observation de la Terre (SPOT) is a French commercial optical multispectral satellite launched in 1986 and initially provided spectral bands at 20 m spatial resolution. The latest missions SPOT-6 and SPOT-7 were launched in 2012 and 2014, respectively. Both of these satellites provide data over five bands (one panchromatic and four multispectral bands at 1.5 m and 6 m spatial resolutions, respectively). The satellite has a temporal resolution of a single day. Other than this, several other very high-resolution multispectral data providing commercial satellites available are WorldView, Earth Resource Observation Satellite (EROS), RapidEye, QuickBird, IKONOS, etc. (Giri et al., 2013).

Sentinel-2 is a newly introduced mission from European Space Agency (ESA) to collect optical imagery around the whole globe at a high spatial resolution of 10-60 m. It collects multispectral data over 13 regions of the electromagnetic spectrum including three new bands on vegetation red-edge region (desirable to detect vegetation changes). The system contains two twin satellites (Sentinel-2A launched in 2015, and Sentinel-2B launched in 2017) which help reduce the revisit time of the satellites to less than five days. The availability of Sentinel-2A and

2B data is viewed as an invaluable resource over other available remote sensing data sources due to its high resolution (spectral, spatial, and temporal) and free access (Grabska et al., 2019).

In addition to optical sensors, synthetic aperture radar (SAR) active remote sensing system that employ the microwave spectral region have been identified as useful tools in vegetation studies (see Section 1.4). The data acquisition by SAR sensors is performed at different polarizations. Polarization is simply the geometric orientation of the electric field when an electromagnetic wave such as light propagates. If the electric field is oriented vertically perpendicular to the direction of wave propagation, it is said to be vertically (V) polarized. If it is oriented horizontally perpendicular, then it is horizontally (H) polarized (Damask, 2004).

The Sentinel-1 is an active remote sensing satellite platform comprising of a set of two satellites, Sentinel-1A, and 1B were launched in 2014 and 2016, respectively and are continuously providing data at a spatial resolution down to 5 m and temporal resolution of less than 12 days (Barra et al., 2017). They are different than optical sensors like multispectral instrument of Sentinel-2 in the sense that they collect data throughout 24 hours, all-weather operational, and provide C-band (long wavelength of 5-8 cm) SAR data acquired in dual polarization mode (VH and VV amplitude bands) (Torres et al., 2012).

1.4 Spectral Information for Sensing Vegetation Properties

Spectral bands are the individual layers of data provided by the sensors over various regions of the electromagnetic spectrum. The number of spectral bands provided by different sensors differs from each other and can cover different ranges. Spectral vegetation indices (SVIs) are developed from mathematical operations between the available spectral bands of optical multispectral or hyperspectral data based on the relationships between different spectral bands. SVIs have been extensively used for various vegetation biophysical characteristics estimation

such as the amount of chlorophyll and water content and vegetation vigor (Jones and Vaughan, 2010). The sensors onboard Landsat and MODIS satellites with fewer spectral bands are applicable for the derivation of traditional but widely used SVIs such as Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and Normalized Vegetation Infrared Indices (NDIIs) (Townsend et al., 2012; Rahimzadeh-Bajgiran et al., 2018). The shortcomings of these traditional SVIs to monitor subtle changes in tree canopy can be overcome using Sentinel-2 images which allow us to calculate a number of red-edge SVIs (sensitive to minute vegetation changes) like Chlorophyll Red-Edge (ClRE), Inverted Red Edge Chlorophyll Index (IRECI), Modified Simple Ratio (MSR), Red-Edge NDVI, and many more (Majasalmi and Rautiainen, 2016; Hawrylo et al., 2018) as will be discussed in Chapters 2 and 3. The numerical values of SVIs vary, but for common normalized SVIs like NDVI and EVI range from -1 to +1. For NDVI, the index value of one is interpreted as a high vegetation amount while values under 0.1 are interpreted as non-vegetated features (Weier and Herring, 2000).

The Sentinel-1 sensor sends the C-band signals to the earth surface and records the signal backscatter. Generally, there are five band frequencies, namely i) P-band, ii) L-band, iii) S-band, iv) C-band, and v) X-band where SAR operates and have been widely used in vegetation mapping. P, L, and S bands are greater in wavelength (~68 cm, ~23.9 cm, and ~9.1 cm respectively) than that of C-band (~5.6 cm) while X-band is even shorter in wavelength (~3.1 cm) (Van Beijma et al., 2014; Main et al., 2016). The radar of Sentinel-1 is capable of transmitting the signals in both horizontal and vertical polarizations and then receiving them in both of those polarizations (European Space Agency, 2016). As SAR waves can penetrate inside canopy, they are capable of providing vegetation structural information and can improve the

estimation of vegetation parameters such as leaf area index and biomass when incorporated with optical data (Sano et al., 2005).

1.5 Monitoring Pest-induced Defoliation using Optical Remote Sensing

Forest pests control is a critical component of successful forest management (Silva et al., 2013). Ground-based survey of defoliated forest stands is costly and sometimes unaffordable and might be impossible to do it at the landscape scale. In addition, the exact location and damage severity is essential for the preparation of health management plans of the host trees against the pests. Remote sensing techniques can provide invaluable information about forest health condition to the forest managers (Silva et al., 2013; Hall et al., 2016). Switching the currently practiced surveying/monitoring methods to remote sensing techniques, information can be summarized for a vast area into a few satellite-derived maps with a high temporal, spatial resolution, making it useful for monitoring landscape level disturbances such as SBW defoliation (Adelabu et al., 2012). Further, the availability of sufficient satellite image acquisition sensors with varying resolutions provides us with multi-scale and multi-stage monitoring of insect-pest borne forest hazards (Silva et al., 2013). The use of ASMs has a long history for the monitoring of pest-induced stresses in North American region, but it demands highly experienced observers and an aircraft from which the observer can record the defoliated areas from non-defoliated ones (Ciesla, 2000). This method is rather costly, human resource intensive and subjective. Also, adverse weather during the time of flight might undermine the quality of data obtained (Hall et al., 2016).

Remote sensing technologies were thought to be very expensive, had technological incapabilities, difficulties in the use and interpretation of data, and were unable to exactly respond to the issues on forest health if we look back to the past endeavors (Peterson et al.,

1999). Recently, a large body of remote sensing sensors and platforms have become available that can help us overcome the shortcomings related to the unavailability of optical remote sensing imagery and their costs (Kuenzer et al., 2014). Remote sensing technologies for defoliation detection have evolved with time from the availability of Landsat series (started from 1972) with coarse temporal resolution to coarse spatial but high temporal resolution MODIS data (satellites launched in 1999 and 2002) to the availability of Sentinel-2 data (satellites launched in 2015 and 2017) which provide both appropriate spatial and temporal resolutions. Together with these events, a large number of high-resolution private satellites like IKONOS (launched in 1999), QuickBird (launched in 2001), and SPOT (started in 1986) were also added.

Eklundh et al. (2009) used MODIS time-series data to map the defoliation caused by pine sawfly (*Neodiprion sertifer*) in Scots pine (*Pinus sylvestris*) stands and found that MODIS data are good for locating the damage but not for its quantification. Similarly, De Beurs and Townsend (2008) used MODIS-based SVIs to estimate the defoliation caused by gypsy moth (*Lymantria dispar* L.) and concluded that the data were useful to monitor defoliation in larger patches of forest (greater than 0.63 km^2). SPOT images also have a long history of use in insect-induced defoliation mapping. Ciesla et al. (1989) used the color composite of SPOT-1 images to map the defoliation caused by gypsy moth in south central Pennsylvania and western Maryland. They attributed gray and black coloration in the image to be defoliated areas. Among all the available sensors, the long history of the Landsat series contributes most in the field of pest damage evaluation (Hall et al., 2006; Townsend et al., 2012; Hall et al., 2016).

Even though a variety of sensors is available, they all have their own advantages and disadvantages. Coarse-resolution sensors like that of MODIS are suitable to be used in large patches of continuous forests and do not perform well in fragmented forests (Eklundh et al.,

2009). Although commercial satellites provide high-resolution data to be used even in fragmented areas, they are very costly and cannot be used for regional level monitoring (Jepsen et al., 2009). As a solution to the above-mentioned issues, Landsat series with a spatial resolution of 30 m and a swath width of 185 km are viewed as a data source for stand-level pest monitoring but again the temporal resolution of 16 days will be a hindrance to accurately detect pest induced changes over a tree canopy (Hall et al., 2016). Therefore, Sentinel-2 datasets (Sentinel-2A and -2B) with a temporal resolution of fewer than five days, the spatial resolution of 10-60 m, a swath width of 290 km, and availability of three additional bands in the red-edge position offer desired improvements over Landsat series for insect and pest-related defoliation monitoring. The presence of three bands in the red-edge region of electromagnetic spectrum provided by Sentinel-2 data will be useful for the detection of the stress-induced spectral signature shift in vegetation and help us with early detection of forest damage (Silva et al., 2013).

There is no universal method for detecting defoliation in the tree using remote sensing techniques (Tewkesbury et al., 2015). Both single-date and multiple-date images could be used for forest defoliation detection but the use of multiple-date imagery is more reliable (Hall et al., 2016). A review of remote sensing techniques applied for change detection caused by insects by Hall et al. (2016) summarized current methods for defoliation detection into five methods: visual analysis, image algebra, classification, spectral mixture analysis, and time series analysis, among which image algebra (Hall et al., 2006) is the one used widely. Townsend et al. (2012) also used the difference between two images (pre-outbreak and post-outbreak) from different dates to quantify gypsy moth-induced defoliation.

In the discipline of forest health and specifically defoliation detection, research on SBW defoliation monitoring utilizing images from space-borne satellites are very limited and to date,

no research has been conducted using Sentinel-2 imagery. However, the capability of Sentinel-2 images in monitoring forest health have been suggested in several other studies (Brovkina et al., 2017; Hawryło et al., 2018; Eschen et al., 2019). Rahimzadeh-Bajgiran et al. (2018) utilized Landsat imagery and SVIs for the successful detection and severity classification of annual SBW defoliation in Quebec and Maine. They further suggested the use of high-resolution Sentinel-2 imagery for this purpose. Fan (2006) also reported the usefulness of multi-temporal Landsat images in the classification of cumulative SBW defoliation. Other than multispectral data, Huang (2015) used hyperspectral imagery (Hyperion data) to map the SBW defoliation in Quebec.

1.6 Host Species Mapping using Optical and SAR Remote Sensing

The application of remote sensing techniques for forest composition mapping has attracted a lot of attention in the research community. Research on species composition is emerging more after the launch of Sentinel-2 satellites, which provide finer spectral and spatial resolutions that were limiting other open access data such as those from Landsat and MODIS (Puletti et al., 2018). White et al. (1995) carried out an unsupervised classification of forest at Lassen Volcanic National Park in California into four classes (pine, fir, non-forest, and non-vegetation) utilizing Landsat TM data with an accuracy of 73%. Species classification studies are also performed using high-resolution data from commercial satellites such as WorldView (Immitzer et al., 2012), RapidEye (Adelabu et al., 2013) and IKONOS (Carleer and Wolff, 2004).

SBW is strictly host specific and attack balsam fir and spruce species. Hence, the mapping of these species is vital to assess the vulnerability of the forest towards SBW attack. Among the few studies of SBW host species mapping using remote sensing, Morris and Bishop (1951) suggested the use of aerial photographs and ground data in combination for rapid rating

of forest vulnerability to SBW damage. Wolter et al. (2008) also utilized remote sensing techniques (Landsat imagery and derived SVIs) to map the distribution and abundance of SBW host species in northern Minnesota and Ontario. They used multi-temporal and multi-seasonal Landsat images for landscape-level outbreak modelling and found it to be effective even in the spatially heterogeneous forests. This study utilized a partial least square regression method to develop a model for host species mapping using 128 ground truth sample plots (basal area by species measured during the period of 2003-2004) and Landsat images (2000-2003). The model built was finally employed to predict the basal area for four forest components which were fir, spruce, deciduous and coniferous trees with R^2 values of 0.64, 0.88, 0.86, and 0.86, respectively.

Since Sentinel-2 data are comparatively newer, there has not been any study conducted on SBW host species mapping using these data; however, there is an abundance of studies using Sentinel-2 for tree species and agricultural crops classification (Immitzer et al., 2016; Karasiak et al., 2017; Nelson, 2017; Puletti et al., 2017; Persson et al., 2018; Grabska et al., 2019). Immitzer et al. (2016) obtained satisfactory results for tree species classification in central Europe with single-date Sentinel-2 imagery, but the use of multi-temporal Sentinel-2 images was suggested to perform better for discriminating the forest species in a heterogeneous setting to capture the phenological differences between different tree species which is a key for species discrimination (Puletti et al., 2017; Persson et al., 2018; Grabska et al., 2019). Further, the combination of SVIs with spectral bands was found to be very useful for distinguishing mixed forests from pure forests, and the model performance during species classification increased significantly with the inclusion of SVIs. Therefore it is highly recommended to use spectral bands in combination with SVIs over the use of individual bands alone (Puletti et al., 2017; Erinjery et al., 2018).

SAR images were previously used mainly for land-use classification (Kurosu et al., 1999; Kurosu et al., 2001; Li and Yeh, 2004). Researchers have shown interest and obtained promising results in using this technology in combination with multi-spectral imagery for the classification of individual tree species (Yu et al., 2018), vegetation types (Sano et al., 2005; Eringery et al., 2018), monitoring of invasive species (Rajah et al., 2019), and habitat assessment (Schmidt et al., 2018). Erinjery et al. (2018) used the combination of Sentinel-2 and Sentinel-1 SAR data to map the tropical rainforest of the Western Ghats and obtained an accuracy of up to 82% with a Random Forest (RF) model. Similarly, Sano et al. (2005) used optical imagery (Landsat-5 SVIs) with SAR (Japanese Earth Resources Satellite-1 (JERS-1)) images and attained an encouraging result with an accuracy of 84% while discriminating the Brazilian savanna vegetation. Schmidt et al. (2018) used the combination of Sentinel-2 optical imagery and Sentinel-1 SAR imagery to assess the habitat quality by mapping vegetation cover in Germany with an overall accuracy of up to 76% using Support Vector Machine (SVM). The increasing trend and success of using SAR imagery combined with optical imagery for vegetation studies inspired its application in the current study.

1.7 Research Problem

As reviewed in the literature, the occurrence of SBW is a cyclic phenomenon. History suggests periodic recurring outbreaks and resulting massive losses in the northeastern forests of America. The current practice of SBW defoliation monitoring using ASMs is not only costly but also time-consuming, human resource-intensive, and subjective. Remote sensing technology can significantly enhance or supplement ASM data for defoliation detection and severity mapping. However, Landsat data with the absence of red-edge bands are restricted to the derivation of very few conventional SVIs only and cannot provide us with the new pool of SVIs which are more

sensitive to slight changes in leaf pigment content, water content and other bio-physical parameters. Also, Landsat series have the temporal resolution of 16 days which makes the regular monitoring of forest health difficult. Therefore the use of new satellite data (Sentinel-2) is expected to bring better results in the field of SBW defoliation monitoring with higher spatial, spectral and temporal resolution than what can be obtained from Landsat. It is expected that light defoliation can be better detected using Sentinel-2 imagery.

The occurrence of SBW outbreak and host species availability are interdependent upon each other, and it is essential to keep track of the abundance of host species to control SBW attack. It was challenging to map individual tree species with the other open-source satellite data such as MODIS and Landsat, and high-resolution commercial satellite data were costly and thus limited only to the small patches of forests. With the limited research in SBW host-species mapping and the unavailability of reliable host abundance information (with the prevalence of SBW attacks), it is challenging to predict the occurrence of the attack and act ahead of time to minimize the loss. Since the launch of Sentinel-2 satellites, researchers have been provided with a great opportunity to map tree species more effectively (having access to finer spatial and spectral information). In addition, the high temporal resolution of less than five days of Sentinel-2 data ensures the availability of a greater number of cloud-free image time-series which is essential to capture the phenology differences between the tree species.

1.8 Research Objectives

To address the problems mentioned above, the specific objectives of our study are:

- ❑ SBW defoliation detection and severity classification:
    - To assess the potential of Sentinel-2 derived SVIs, and spectral bands, especially red-edge bands and SVIs for light SBW defoliation detection and severity classification.

- To investigate the usefulness of multi-year images over single-year imagery for both defoliation detection and host species classification.
- To examine the usefulness of site variables (elevation, slope, aspect) for defoliation detection
- To evaluate the performance of different non-parametric algorithms for SBW defoliation modeling and create the best model

❑ SBW host species classification:

- To investigate the effectiveness of Sentinel-1 and Sentinel-2 satellite imagery alone or in combination for SBW host species classification.
- To examine the usefulness of site variables (elevation, slope, aspect, soil types, TWI, projected climate site index for year 2030 and iBGI) for host species classification.
- To create a model for SBW host species mapping

1.9 Thesis Structure and Overview of Each Chapter

This thesis is divided into four chapters, as shown in Figure 1.2. Chapter 1 provides an introduction into the background of SBW and its outbreaks, traditional and remote sensing techniques for pest-induced defoliation detection with the focus on SBW defoliation detection, the importance of mapping SBW host species and the application of remote sensing for tree host species mapping. The presented information in this chapter provides a backgrounds of the research and methods used in Chapters 2 and 3. Chapters 2 and 3 are written in the format of a manuscript and could be read independently from other chapters. Due to this reason, there could be some information overlap among the chapters. Chapter 1 also summarizes the overall research problems and thesis objectives.

Chapter 2 focuses on the assessment of the capabilities of different Sentinel-2 based SVIs, single spectral bands and available site variables to build a robust model using RF and SVM algorithms for the detection and classification of SBW defoliation. The chapter deals with two approaches for defoliation detection and classification, namely single-year and multiple-year defoliation detection/classification.

Chapter 3 investigates the potential of both Sentinel-1 and Sentinel-2 variables in discriminating the tree host species of SBW. Different models were created for Sentinel-1 variables plus site variables and Sentinel-2 variables plus site variables. Also, all these variables were combined to evaluate their performance in a single integrated model.

Finally, the last chapter integrates Chapters 2 and 3 and concludes the entire research results of this thesis. This chapter seeks to examine the usefulness of integration of defoliation detection and host species distribution results to monitor the SBW defoliation better and mitigate probable losses from the outbreak.

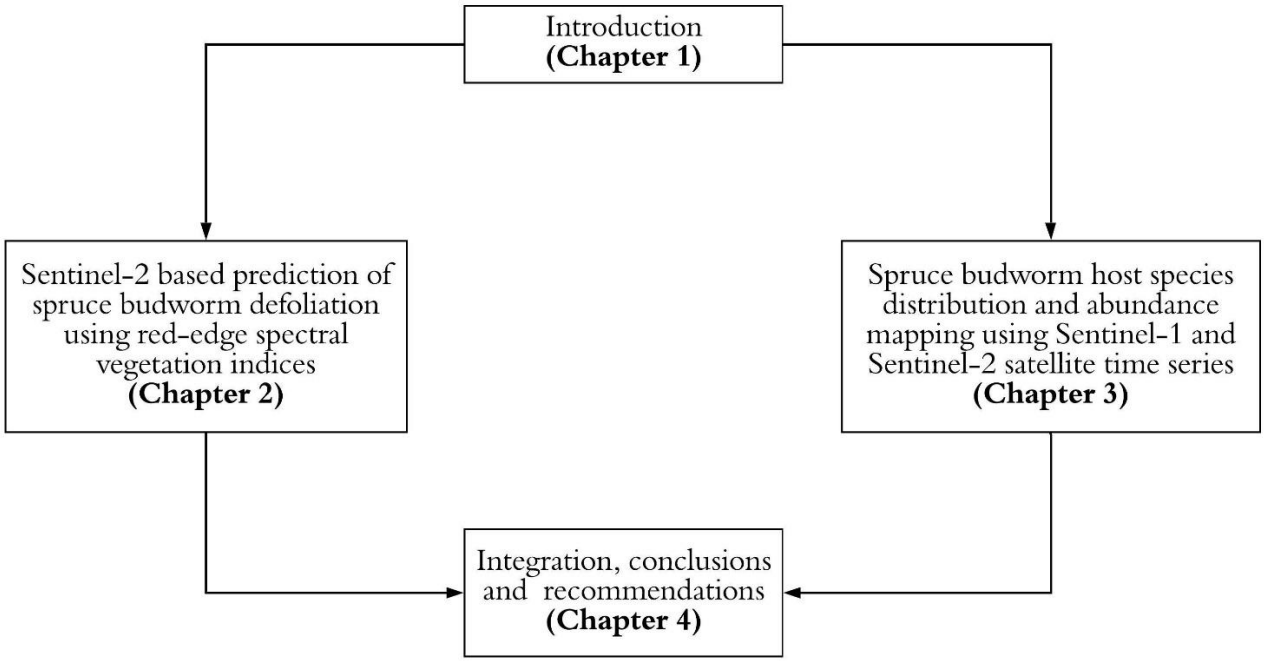


Figure 1.2. The structure of this thesis

CHAPTER 2

SENTINEL-2 BASED PREDICTION OF SPRUCE BUDWORM DEFOLIATION USING RED-EDGE SPECTRAL VEGETATION INDICES

2.1 Introduction

Insects and diseases are the most extensive disturbance agents in forests of Canada and the U.S., affecting millions of hectares per year (Tkacz et al., 2007). Insect and disease damage can cause defoliation at different scales from individual trees to landscapes, leading to mass mortalities depending on the type of host trees and their distribution (Brovkina et al., 2017). Accurate data on forest pest-induced damage extent and severity is essential for pest control and early intervention and consequently host tree protection. This information is also needed for pest risk prediction, a better estimation of economic impacts of outbreaks, quantifying wood supply losses, identifying changes in wildlife habitat, and revising strategic forest management plans.

The eastern spruce budworm (*Choristoneura fumiferana* Clem.; SBW) is a native pest species of North America that causes widespread defoliation of balsam fir (*Abies balsamea* [L.] Mill.) and spruce (*Picea* spp.) trees (MacLean 1984). The pest erupts periodically approximately every three decades in spruce-fir forests, resulting in defoliation of millions of hectares (MacLean, 1980; MacLean and MacKinnon, 1996). The most recent severe eastern SBW outbreak began in Quebec in 2006 which defoliated over nine million ha of forestland by the summer of 2019 (Ministère des Forêts de la Faune et des Parcs., 2019) and is gradually moving towards New Brunswick (NB) and the state of Maine in the U.S.A. (MacLean et al., 2019).

The mortality and severe growth reductions resulting from SBW outbreaks have motivated researchers to track and detect SBW defoliation and monitor its impacts on forest ecosystems (MacLean, 1984). Rather than continuing on the traditional reactive foliage protection using insecticides approach to keep trees alive, forest managers are currently testing

early intervention techniques to detect and suppress the SBW hot spots before substantial defoliation occurs (MacLean et al., 2019). For the early intervention technique, intensive and accurate landscape level monitoring of the forest and early detection of low-level SBW defoliation is desired but has been challenging to achieve (Rahimzadeh-Bajgiran et al., 2018).

SBW larvae strongly prefer to feed on the current-year foliage of host trees but it is the repeated defoliation over several years that kills trees (MacLean, 1980). Therefore defoliation can be classified as both current year/annual defoliation (on the newest age-class of foliage) and cumulative defoliation on all age-classes of foliage. The focus of our study was on current year defoliation, which is applicable to an early intervention approach to control SBW related forest losses and can also serve as an input into decision support systems (MacLean et al., 2019). Generally the estimation of current year defoliation on a landscape level is performed using aerial surveys, with accuracies dependent on the observers' ability to recognize the defoliation severity and identify the location, and also influenced by weather conditions. Also the accuracy of aerial surveys of SBW defoliation is subjective and it is difficult to discriminate nil (0-10%) from low (11-30%) defoliation classes (MacLean and Mackinnon, 1996).

Since the year 2000, many remote sensing sensors and platforms have become available that can be used for pest management applications. Compared to aerial surveys, detection using satellite imagery may be less costly, less subjective and less laborious (Hall et al., 2016; Rahimzadeh-Bajgiran et al., 2018). Although the capability of the newly launched Sentinel-2A and -2B satellite imagery in monitoring forest health has been demonstrated in recent studies (Hawryło et al., 2018) for the detection of Scots pine (*Pinus sylvestris* L.) defoliation, no work has been published dealing with the applications of this sensor for SBW annual defoliation detection. Given the technical advantages of Sentinel-2 satellites over Landsat (higher spectral,

spatial and temporal resolution) they are expected to be more effective for mapping light defoliation. The objective of this research was to evaluate the performance of red-edge spectral vegetation indices (SVIs) derived from Sentinel-2 imagery for the detection and classification of lightly defoliated spruce-fir forests of northern NB.

2.2 Materials and Methods

2.2.1 Study Area and Field Data

The study area is located in northern NB, Canada, where SBW defoliation started in 2014 (MacLean et al., 2019). The forest consists of the elements of both northern boreal coniferous forests and deciduous forests from the south. The forest experiences a milder winter and cooler summer, although NB is a maritime province (Simmons et al., 1984). The primary host species found in the study area are white spruce (*Picea glauca* (Moench) Voss), black spruce (*Picea mariana* [Mill] BSP), red spruce (*Picea rubens* Sarg.) and balsam fir (*Abies balsamea* [L.] Mill.). Figure 2.1 shows the location of the study area in NB, Canada.

Field data were collected during the SBW defoliation surveys of New Brunswick Department of Energy and Resource Development in 2018. The data were from a road-side survey, where each sample point was a sample of one mid-crown branch which are later assigned into broad defoliation classes. A total of 236 sample points divided into five defoliation classes (1=no-defoliation, 2= (1-5) % defoliation, 3= (6-30) % defoliation, 4= (31-70) % defoliation and 5= >70% defoliation) were used for model training (2/3rd of samples) and validation (1/3rd of samples). Most of the samples were in Class 2 and 3 (132 and 54 samples, respectively) as the defoliation level was light in general. Classes 1, 4 and 5 comprised of three, 40, and seven samples, respectively.

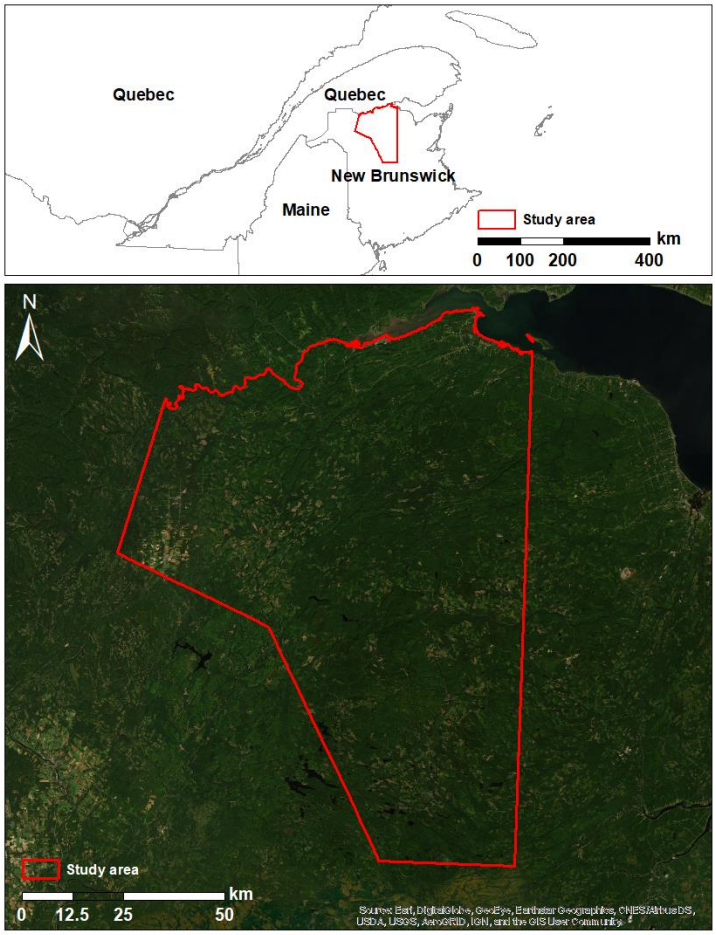


Figure 2.1. The extent of the study area in NB, Canada.

2.2.2 Remote Sensing Data Selection, and Pre-processing

To detect SBW annual defoliation using remotely sensed imagery, information on SBW phenology is essential. SBW has six larval instars (L1 through L6) before it pupates. During feeding, SBW larvae forms a web-like feeding tunnel around several terminal twigs and partially consumed needles become entrapped in this webbing, then dry, and turn a reddish color (Leckie et al., 1992). In 2018 in our study area, peak SBW L5 occurred at mid-June, peak L6 occurred in late June, peak pupae occurred around July10, and peak redness of the forest due to defoliation occurred between peak L6 and peak pupae. L5 and L6 instars of SBW consume about 95% of the host needles while other larval instars consume less than 5% altogether (Miller, 1977).

Therefore, six images dated June 26 and July 21 2018 and July 1st 2016 (two images per date) were selected to represent before and after defoliation conditions in 2018 and conditions two years before. Cloud-free Sentinel-2 Level 1C Top-Of-Atmosphere reflectance products (L1C-TOA) were collected from https://earthexplorer.usgs.gov/. Using Sentinel-2 SNAP toolbox the L1C-TOA imagery were converted to L2A surface reflectance products.

2.2.3 Defoliation Detection Approaches

SVIs have been used in previous remote sensing research to detect and estimate pest-induced defoliation and damage using Landsat and MODIS imagery (Townsend et al., 2012; Hall et al., 2016). Recently the use of a combination of SVIs for SBW defoliation detection was suggested by Rahimzadeh-Bajgiran et al. (2018). Additionally, the potential use of red-edge SVIs has been suggested instead of traditional SVIs for pest-induced defoliation detection (Hall et al., 2016). Sentinel-2 data enable mapping defoliation using similar spectral bands to those of Landsat; however, they have several advantages including i) higher spatial (10-20 m) resolution, ii) greater temporal (at least five days with A and B satellites combined) resolution, and iii) three additional spectral bands in the red-edge region. A total of 29 predictor variables were used to build detection and classification models using Random Forest (RF) and Support Vector Machine (SVM) algorithms with nine single spectral bands and 20 SVIs using L2A products at 20 m resolution (Table 2.1). In addition, elevation data were also used as a site variable for modeling based on Chen et al. (2020 in Review) and Chen et al. (2018).

Table 2.1. Sentinel-2 derived predictor variables used for the development of defoliation detection and classification models.

Acronym (Band/SVIs)*	Formula	Reference
b2: Blue (490 nm) b3: Green (560 nm) b4: Red (665 nm) b5: Red-Edge (705nm) b6: Red-Edge (740nm) b7: Red-Edge (783 nm) b8a: NIR (865 nm) b11: SWIR (1610 nm) b12: SWIR (2190 nm)	------	------
WDRVI	((0.01×b7)-b4)/((0.01×b7+b4)+ ((1-0.01)/(1+0.01)))	(Majasalmi and Rautiainen, 2016)
S2REP	705+35× ((((b7+b4)/2)-b5)/(b6-b5))	(Guyot and Baret, 1988)
NDVI	(b8a-b4)/(b8a+b4)	(Rouse et al., 1974)
NDVI45	(b5 - b4)/(b5 + b4)	(Delegido et al., 2011)
NDVI65	(b6 - b5)/(b6 + b5)	(Gitelson and Merzlyak, 1994)
MTCI	(b6-b5)/(b5-b4)	(Dash and Curran, 2007)
MSR	((b7/b4)-1)/√((b7/b4)+1)	(Chen, 1996)
IRECI	(b7-b4) × (b6/b5)	(Clevers et al., 2000)
GNDVI	(b8a-b3)/(b8a+b3)	(Gitelson et al., 1996)
EVI8	2.5× (b8a-b4)/(1+b8a+6×b4-7.5×b2)	(Huete et al., 2002)
EVI7	2.5× (b7-b4)/(1+b7+6×b4-7.5×b2)	(Majasalmi and Rautiainen, 2016)
ClRE	(b7/b5)-1	(Gitelson et al., 2003)
NDII11	(b8a-b11)/(b8a+b11)	(Hardisky et al., 1983)
NDII12	(b8a-b12)/(b8a+b12)	(Key et al., 2002)
ARI1	(1/b3)-(1/b5)	(Gitelson et al., 2001)
ARI2	(b8a/b3)-(b8a/b5)	
CRI1	(1/b2)-(1/b3)	(Gitelson et al., 2002)
CRI2	(1/b2)-(1/b5)	
MCARI	1-((0.2) × (b5-b3)/(b5-b4))	(Hawrylo et al., 2018)
Red-Edge NDVI	(b8a-b6)/(b8a+b6)	(Gitelson and Merzlyak, 1994)

*: NIR: Near Infrared, SWIR: Shortwave Infrared, WDRVI: Wide Dynamic Range Vegetation Index, S2REP: Sentinel-2 Red Edge Position, NDVI: Normalized Difference Vegetation Index, MTCI: MERIS Terrestrial Chlorophyll Index, MSR: Modified Simple Ratio, IRECI: Inverted Red Edge Chlorophyll Index, GNDVI: Green NDVI, EVI: Enhanced Vegetation Index, ClRE: Chlorophyll Red Edge, NDII: Normalized Difference Infrared Index, ARI: Anthocyanin Reflectance Index, CRI: Carotenoid Reflectance Index, MCARI: Modified Chlorophyll Absorption in Reflectance Index

We used RF and SVM as classifiers because of their encouraging results in several remote sensing and insect- forest disturbance studies (Pal, 2005; Hawryło et al., 2018; Rahimzadeh-Bajgiran et al., 2018). Red edge SVIs were used to estimate the chlorophyll content in leaves (Gitelson and Merzlyak, 1994; Rahimzadeh-Bajgiran et al., 2012a), carotenoid and anthocyanin indices to estimate the stress pigments in plants (Gitelson et al., 2002), near-infrared (NIR) indices to measure leaf structure and amount, and short wave infrared (SWIR) indices to assess leaf water content (Townsend et al., 2012; Rahimzadeh-Bajgiran et al., 2018).

The RF is based on an ensemble decision tree (Pal, 2005) while the SVM (Lin, 2004) is based on the principle of maximization of the margin between the classes using training data by finding a hyperplane that separates the groups best from each other. Defoliation can be detected by estimating changes in reflectance in defoliated forest stands compared with their pre-defoliation condition (Townsend et al., 2012; Rahimzadeh-Bajgiran et al., 2018). Defoliation detection for this study was conducted using two approaches: i) evaluating the change occurring in a single year, assuming the forest before the onset of activity of SBW as the healthy state and the forest after being defoliated as the defoliated state (here June 26 2018 as healthy state and July 21 2018 as defoliated); and ii) using the change occurring over multiple years where the year before the outbreak was taken as healthy year and the year after the outbreak was taken as defoliated year (here July 1st 2016 as healthy state and July 21 2018 as defoliated).

2.3 Results

2.3.1 SBW Defoliation Detection and Classification at 20 m using RF and SVM

Figure 2.2 shows the errors for five single-year defoliation detection model formulations using RF to discriminate between defoliated and non-defoliated (less than 5%) states. The five model formulations included all 29 variables, only the 20 SVIs, the best 10 variables, the best 5

variables, and the best five variables plus elevation. The combination of the best five variables (EVI7, MCARI, IRECI, NDII11 and MSR, respectively based on mean decrease in accuracy (MDA) score) reduced the model out-of-bag (OOB) error to 20%, while the inclusion of elevation further decreased the OOB error to 17% (Figure 2.2). The best model formulation did not include any individual Sentinel-2 bands.

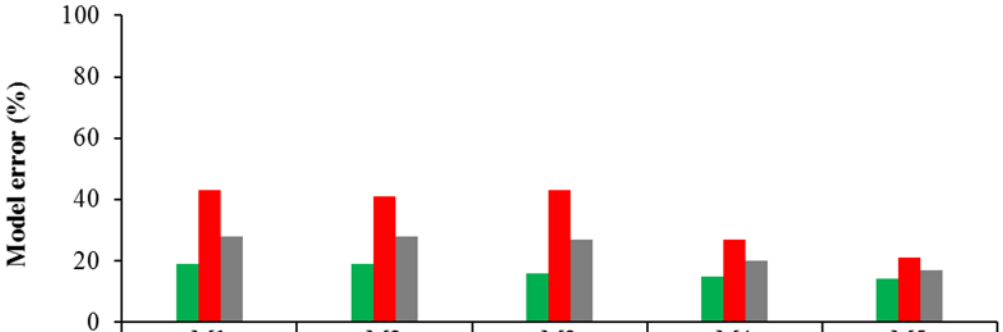
The detected defoliation was then classified into three classes of non-defoliated (0%-5%), light (6%-30%) and moderate (31%-70%) because of the unavailability of sufficient training data for severely defoliated plots. Figure 2.3 shows the performance of five model formulations, where the combination of the best five variables (EVI7, MCARI, IRECI, NDII11 and MSR, respectively) showed the best performance. The OOB error for the best model for classifying defoliation into non-defoliated, light, and moderate classes was 42%, while the inclusion of elevation decreased the OOB error to 32%. Class-wise errors of the model were 16%, 52%, and 39% for no defoliation, light, and moderate classes, respectively after the inclusion of the elevation (Figure 2.3). For both defoliation detection and classification, elevation ranked third most important variable after EVI7 and MCARI based on MDA scores. It should be noted that misclassifications mainly occurred between neighboring classes where light defoliation overlapped medium and no defoliation equally, and medium defoliation was misclassified with light defoliation.

For multiple-year defoliation detection and classification, the best combination was found to be ARI1, NDII11, EVI7, IRECI, MSR plus elevation which resulted in a total OOB error of 21% with an error of 20% for detection of defoliated pixels and 22% for detection of non-defoliated pixels. For classification, the combination of ARI1, NDII11, EVI7, IRECI, MSR and elevation was the best with 33% total OOB error (non-defoliated pixels 19%, light defoliated

pixels 52% and, moderately defoliated pixels 36%). These results were similar to single-year results with the single-year approach slightly outperforming the multiple-year approach.

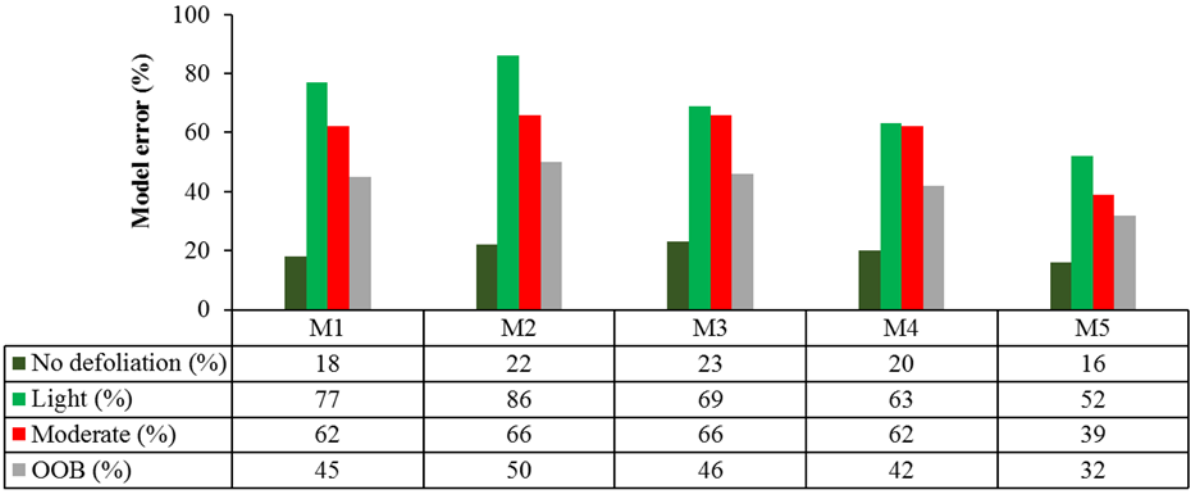
The SVM was also applied for different combinations of predictors similar to Figures 3 and 4 for both defoliation detection, and classification using single- and multiple-year approaches and was found to be slightly less robust than the RF models in detecting and classifying defoliation. The best Sentinel-2 variable combination for single-year defoliation detection was IRECI, ARI1, MCARI, MSR and EVI7, whereas for multiple-year it was NDII11, EVI7, MSR, NDVI45 and Red-Edge NDVI.

Figure 2.4 shows defoliation maps produced using the best RF model and the single-year approach. In the total study area of 475,888 ha, the predicted areas for non-defoliated, light and moderate were 311,182 ha, 26,892 ha, and 137,814 ha, respectively. In agreement with the ground sample points in Figure 2.4c, the amount of defoliation was found to be more in the northern and central parts of the study area.



	M1	M2	M3	M4	M5
No defoliation (%)	19	19	16	15	14
Defoliation (%)	43	41	43	27	21
OOB (%)	28	28	27	20	17

Figure 2.2. Error comparison of the best detection models (RF model at 20 m for single-year defoliation). Here, the band combinations are Model 1 (M1, all 29 variables), M2 (only 20 SVIs), M3 (best 10 variables), M4 (best 5 variables: EVI7, MCARI, IRECI, NDII11 and MSR) and M5 (EVI7, MCARI, IRECI, NDII11, MSR and elevation).



	M1	M2	M3	M4	M5
No defoliation (%)	18	22	23	20	16
Light (%)	77	86	69	63	52
Moderate (%)	62	66	66	62	39
OOB (%)	45	50	46	42	32

Figure 2.3. Comparison of the performance of various variable combinations for defoliation classification for the best models (RF model at 20m for single-year defoliation). Here, the band combinations are Model 1 (M1, all 29 variables), M2 (only 20 SVIs), M3 (best 10 variables), M4 (best 5 variables: IRECI+EVI7+NDVI45+MCARI+MSR), and M5 (IRECI+EVI7+NDVI45+ MCARI+ MSR+ Elevation). Here, the defoliation ranges for "No-defoliation (nil)", "Light" and "Moderate" are 0%-5%, 6%-30% and 31%-70%, respectively.

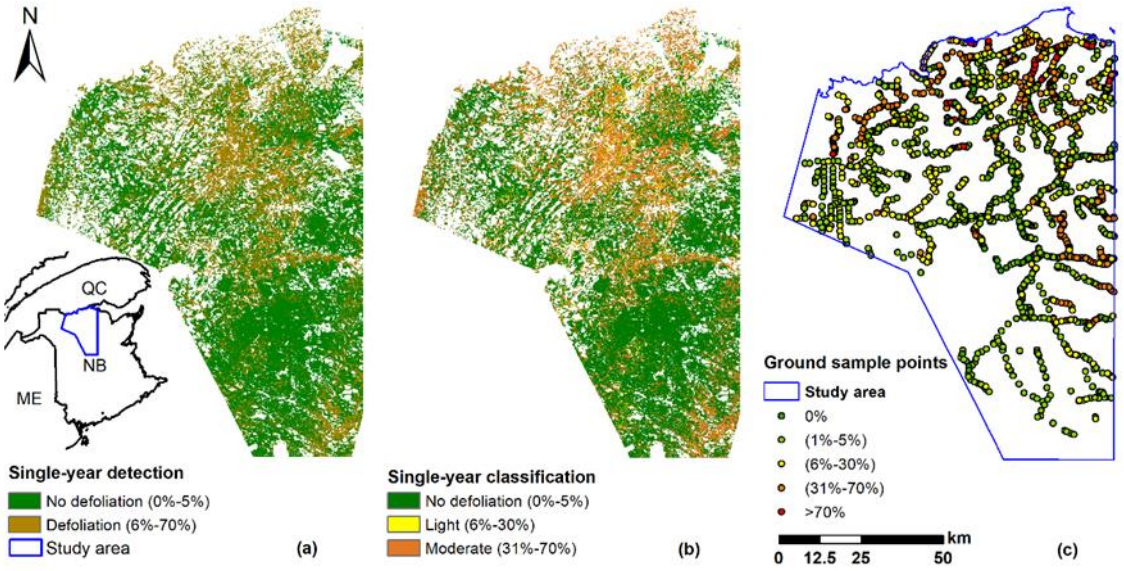


Figure 2.4. (a). Defoliation detection prediction map of only host species using the best RF model for single-year approach at 20 m. (b). Defoliation classification prediction map using RF. (c). Location and severity level of ground truth SBW defoliation.

Figure 2.5 shows the defoliation classification map (single-year 20m) extracted out for mature and immature host species (white spruce, red spruce, black spruce, balsam fir, mixed balsam fir-spruce, and host species mixed with non-host species) for a subset of the study area (118,414 ha) using forest composition data available on GeoNB (http://www.snb.ca/geonb1/e/index-E.asp). According to (MacLean, 1980) higher vulnerability occurs in mature host species while it is comparatively lower in immature ones. The estimated area for no defoliation, light and moderate classes were 15,925 ha, 3,286 ha, and 12,441 ha, respectively for young stands while 44,703 ha, 7,747 ha, and 34,312 ha, respectively for mature stands. The stands are considered mature if older than approximately 40 years. Our observation implied the presence of high defoliation in mature stands which could be supported to some extent by the study of MacLean (1980).

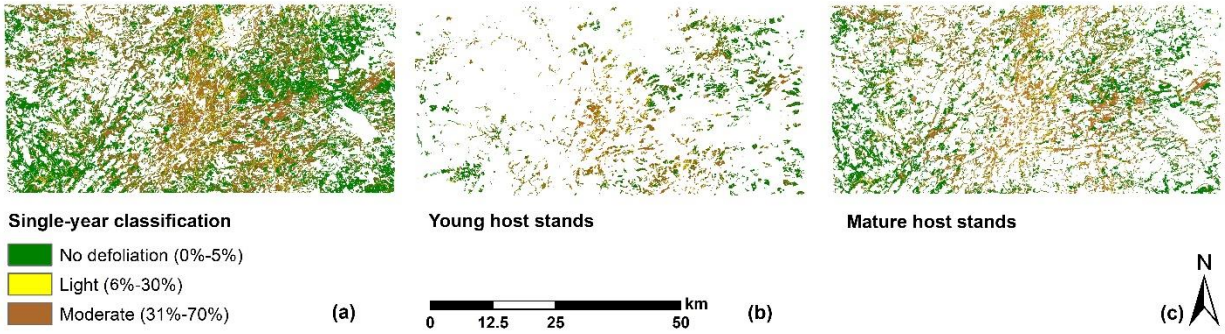


Figure 2.5. SBW defoliation as a function of the maturity of host species using RF single-year model (central coordinates of the image: 47.69°N, 66.74°W).

2.3.2 Model Validation for Single-year and Multiple-year Defoliation Detection and Classification

The post-classification and model validation results for both single- and multiple-year defoliation detection and classification approaches using RF and SVM were calculated, and are presented in Table 2.2 (detection models) and Table 2.3 (classification models). Using the RF

defoliation detection model, the overall accuracy (OA) and kappa coefficient (*K*) were 86% and 0.73 for the single-year approach and, 78% and 0.57 for multiple-year, respectively. Similarly, the SVM model gave overall accuracy and kappa coefficient of 80% and 0.58 for the single-year approach and 78% and 0.55 for multiple-year, respectively. This shows both RF and SVM models and change detection approaches used to detect defoliation had good performances with the single-year approach using RF model outperforming others.

For single-year classification, overall accuracy and kappa coefficient for RF method were 71% and 0.49, respectively and for multiple-year classification, they were 63% (OA) and 0.41 (*K*). For single-year classification, overall accuracy and kappa coefficient for SVM method were 65% and 0.40, respectively and for multiple-year classification, they were 51% (OA) and 0.19 (*K*).

Table 2.2. Confusion matrix and accuracy estimation (%) of the detection models. PA is the producer's accuracy, and UA is the user's accuracy. Bold values are % correct classes.

Detection models					
RF single-year		No Defoliation	Defoliation	PA (%)	UA (%)
	No Defoliation	**84**	10	84	91
	Defoliation	16	**90**	90	82
Overall accuracy:86% and kappa coefficient:0.73					
SVM single-year		No Defoliation	Defoliation	PA (%)	UA (%)
	No Defoliation	**92**	35	92	77
	Defoliation	08	**65**	65	87
Overall accuracy: 80% and kappa coefficient:0.58					
RF multiple-year		No Defoliation	Defoliation	PA (%)	UA (%)
	No Defoliation	**82**	24	82	70
	Defoliation	18	**76**	76	86
Overall accuracy: 78% and kappa coefficient:0.57					
SVM multiple-Year		No Defoliation	Defoliation	PA (%)	UA (%)
	No Defoliation	**71**	16	84	75
	Defoliation	29	**84**	71	81
Overall accuracy: 78% and kappa coefficient:0.55					

Table 2.3. Confusion matrix and accuracy estimation (%) of the classification models. PA is the producer's accuracy, and UA is the user's accuracy. Bold values are % correct classes.

Classification models					
RF single-year	No Defoliation	Light Defoliation	Moderate Defoliation	PA (%)	UA (%)
No Defoliation	**92**	38	17	92	82
Light Defoliation	8	**25**	25	25	29
Moderate Defoliation	0	37	**58**	58	70
Overall accuracy:71% and kappa coefficient:0.49					
SVM single-year	No Defoliation	Light Defoliation	Moderate Defoliation	PA (%)	UA (%)
No Defoliation	**90**	30	11	90	83
Light Defoliation	0	**10**	22	10	33
Moderate Defoliation	10	60	**67**	67	43
Overall accuracy:65% and kappa coefficient:0.40					
RF multiple-year	No Defoliation	Light Defoliation	Moderate Defoliation	PA (%)	UA (%)
No Defoliation	**80**	20	57	80	71
Light Defoliation	16	**50**	14	50	67
Moderate Defoliation	4	30	**29**	29	22
Overall accuracy:63% and kappa coefficient:0.41					
SVM-multiple-year	No Defoliation	Light Defoliation	Moderate Defoliation	PA (%)	UA (%)
No Defoliation	**88**	71	17	88	56
Light Defoliation	8	**15**	25	25	63
Moderate Defoliation	4	14	**58**	14	20
Overall accuracy:51% and kappa coefficient:0.19					

2.4 Discussion

Sentinel-2 data were evaluated for SBW defoliation detection and classification in a lightly defoliated study area. We found that the suggested model can effectively detect light defoliation but classification was not as accurate as detection. Among SVIs, except NDII, which is known to be a good index for pest induced defoliation detection (Townsend et al., 2012), all of the best indices were red-edge SVIs suggesting that these SVIs may generally outperform

traditional ones such as NDVI and EVI. Among the best SVIs identified in this study, the performances of EVI and NDII11 are in agreement with previous results (Rahimzadeh-Bajgiran et al., 2018) where the best SVIs used to detect and classify SBW defoliation in Quebec using Landsat 5 imagery were EVI, NDVI, and NDMI. However, in this study, EVI7 (red-edge EVI) outperformed EVI8 used in (Rahimzadeh-Bajgiran et al., 2018). We also found the anthocyanin based index ARI1 a robust index to detect SBW defoliation which is also in accordance with the findings of a research estimating the defoliation of Scots pine (*Pinus sylvestris* L.) stands in Poland (Hawryło et al., 2018).

Several hyperspectral as well as multispectral image based studies on forest health decline have found that red-edge indices are very sensitive to even a small amount of chlorophyll content change in the leaf surface of mixed coniferous forest (Hall et al., 2016) as well as spruce decline (Brovkina et al., 2017). Our research also showed that the majority of the important vegetation indices to detect defoliation (IRECI, EVI7, MCARI, ARI1, and MSR) are red-edge vegetation indices. These results are similar to those of Clevers and Kooistra (2011). Several simulation studies have also been conducted to show the importance of the red edge region of Sentinel-2 in determining the chlorophyll content of the canopy (Delegido et al., 2011; Majasalmi and Rautiainen, 2016) and found that these SVIs are robust for estimating canopy chlorophyll level. These results could be used to justify our findings by relating the defoliation and canopy to chlorophyll decline. In all of the single-year and multiple-year defoliation detection and classification approaches in our study, MSR and IRECI played a pivotal role. Literature also mentions IRECI is the best index for the estimation of canopy chlorophyll and MSR is one of the strongest indices to determine leaf area index (Frampton et al., 2013). Although some studies have indicated better performance of GNDVI as compared to NDVI

(Gitelson et al., 1996; Hawryło et al., 2018) this was not observed in our study as well as by Rahimzadeh-Bajgiran et al. (2018).

For SBW defoliation classification, the inclusion of elevation in the model increased the accuracy for non-defoliated class by 4%, light defoliated class by 11%, and moderately defoliated class by 23%. The importance of elevation information in SBW defoliation observed here is corroborated by previous studies (Chen et al., 2018; Chen et al., 2020 in review).

Comparing our results with the results from aerial sketch mapping of MacLean and Mackinnon (1996), the similar issue for the difficulty in distinguishing between non-defoliated and light defoliated has been seen here. Still, the per cent defoliation range we used for no defoliation is rather narrower (0%-5%) and the combination of non-defoliated with defoliated classes into a single class gave us an overall accuracy of 86%.

The comparison of accuracies between different machine learning models for the detection and classification of defoliated pixels showed both machine learning models used (RF and SVM) were generally similar in terms of performance, as also found by (Pal, 2005) and (Hawryło et al., 2018) but RF performed slightly better than SVM in our case. In other fields, RF has often been found to outperform SVM when it comes to multi-class classification (Statnikov et al., 2008).

2.5 Conclusion

As a solution for the drawbacks of previous studies where satellite (e.g. Landsat) spatial and spectral resolutions were the main limitations for light defoliation detection over dense forest cover, Sentinel-2 imagery was evaluated and found to be effective using the suggested variables and models. We found that the classification of light defoliation is still challenging mainly because of the differentiation of lightly defoliated pixels from non-defoliated ones at lower

defoliation levels. From a practical standpoint, current year SBW defoliation classification is the most valuable information for forest pest early intervention strategies. Any remote sensing studies attempting to estimate current year defoliation that are not timed precisely with the approximate 2-week red stage SBW defoliation are actually assessing cumulative multi-year defoliation which is of less significance. In our study, we used the exact time span for defoliation assessment leading to high detection accuracies but comparatively lower classification accuracies. Given the scattered nature of early defoliation stages in the study area and lack of landscape-scale field and aerial data, Sentinel-2 red edge SVIs can be beneficial to map light and scattered defoliation and provide a basis for subsequent year evaluations. We expect that our suggested method could be applicable for other pest- or pathogen-induced defoliation assessment where early stage distribution of affected trees is of interest.

CHAPTER 3

SPRUCE BUDWORM HOST SPECIES DISTRIBUTION AND ABUNDANCE MAPPING USING SENTINEL-1 AND SENTINEL-2 SATELLITE TIME SERIES

3.1 Introduction

Spruce budworm (*Choristoneura fumiferana* Clem; SBW) has been the primary forest disturbance agent in Northeastern forests of the United States and Canada over the past centuries (Kettela, 1983; MacLean et al., 2019). The last SBW outbreak in these regions from 1967 to 1993 caused the defoliation of over 50 million ha of forests (Kettela, 1983; MacLean et al., 2019) with high rates of tree growth losses and increased mortality (Chen et al, 2017a, b). A new SBW outbreak began in 2006 in Quebec and in 2015 in New Brunswick (NB) in Canada, and is currently moving toward neighboring locations in Canada and the U.S.A.

This host-specific pest selectively attacks spruce (*Picea* spp.) and balsam fir (*Abies balsamea* (L.) Mill.) and therefore its damage extent and severity is primarily attributed to the abundance and distribution of these species (Blais, 1983; MacLean and Piene, 1995; Wolter et al., 2008). Susceptibility (amount of defoliation) among host species is highest for balsam fir, followed by white spruce (*Picea glauca* (Moench) Voss), red spruce (*Picea rubens* Sarg.), and black spruce (*Picea mariana* (Mill.) (Hennigar et al., 2008). Vulnerability (severity of growth reduction and mortality level) is highest for mature balsam fir followed by mature spruce, immature balsam fir and immature spruce (MacLean, 1980). Mortality of forest stands during SBW attack is highly influenced by species composition (hosts and non-hosts availability) and defoliation level of balsam fir in mixedwood and hardwood stands is less than in pure host stands (Su et al. 1996; Zhang et al., 2018). The widespread, periodic outbreaks of this pest not only differentially kill the individual species, but also alter the structure and function of forests (Fleming et al., 2002). Up-to-date and accurate tree host species maps are required for

monitoring and any early intervention activities (MacLean et al., 2019; Johns et al., 2020) and in addition they provide important insights into the patterns, cycle, severity, and impact of large-scale insect disturbances like SBW (Bouchard et al., 2006; Wolter and Townsend, 2011).

Over the past few decades, remote sensing techniques have been flourishing in forest composition mapping using different platforms including optical multispectral satellite imagery such as Landsat (Schmitt et al., 1996; Wolter et al., 2008; Thapa et al., 2020) and Sentinel-2 (Immitzer et al., 2016; Grabska et al., 2019). Most studies on tree species classification using satellite multispectral imagery have been conducted using individual spectral bands (Immitzer et al., 2016; Persson et al., 2018; Grabska et al., 2019) and few have combined spectral vegetation indices (SVIs) and spectral bands (Wolter et al., 2008; Puletti et al., 2017). Results from previous research show that the inclusion of SVIs improved the performance of the classification over the use of spectral bands alone (Wolter et al., 2008; Puletti et al., 2017). Tree species classification using Sentinel-2 imagery has resulted in promising outcomes from single-date images (Immitzer et al., 2016), however, classification using multi-temporal imagery (time series) generally outperformed those using single-date data (Persson et al., 2018; Grabska et al., 2019).

Several studies have used synthetic aperture radar (SAR) and optical multispectral imagery in combination for invasive species monitoring (Rajah et al., 2019), vegetation type classification (Sano et al., 2005; Eringery et al., 2018), individual tree classification (Yu et al., 2018), and habitat assessment (Schmidt et al., 2018). SAR images alone have also been used for forest classification (Ranson and Sun, 1994; Rignot et al., 1994). Most studies found improved classification results when SAR was combined with optical multispectral imagery (Sano et al., 2005; Wolter and Townsend, 2011; Yu et al., 2018; Erinjery et al., 2018; Schmidt et al., 2018).

Studies on SBW host species mapping using remote sensing data are limited (Wolter et al., 2008; Legaard et al., 2020). In addition, review of the literature indicated that a combination of Sentinel-1 SAR and Sentinel-2 multispectral imagery has mainly been used for vegetation cover type mapping rather than forest composition classification and studies generally used single spectral bands and traditional SVIs such as Normalized Difference Vegetation Index (NDVI) (Immitzer et al., 2016; Eringery et al., 2018). The technical advantages of recently-launched Sentinel-2A and -2B satellites over similar non-commercial satellites like Landsat (higher spectral, spatial and temporal resolutions and additional spectral bands in the red-edge regions), make them attractive for forest composition mapping. Sentinel-1 SAR images not only can provide spectral information from a different perspective, but also may serve as an alternative for the cloudy or unavailable images from Sentinel-2.

In this study, we used the combined time-series of Sentinel-1 SAR, several Sentinel-2 spectral bands and SVIs in particular red-edge and narrow-band SVIs, as well as site variables as a novel approach to map SBW host and non-host tree species in northern NB, Canada. The specific objectives were 1) to evaluate the capability of Sentinel-1 SAR and Sentinel-2 optical multispectral imagery alone or in combination for SBW host species classification, 2) to investigate the effectiveness of the inclusion of site variables with satellite imagery for SBW host species mapping, and 3) to produce fine-scale maps of SBW host species abundance that show their current spatial distributions.

3.2 Materials and Methods

3.2.1 Study Area

The study area is located in the northern part of NB in Canada (Figure 3.1). The study area consists of the boreal forests in the northern part and Canadian Acadian forests in the south

consisting of the elements of both of the forest types. The public forest of NB consists of a variety of hardwood and softwood species, but the balsam fir and spruce (SBW host species) comprise 55% of the forest volume (Erdle and Ward, 2008). Tree species in the study area include balsam fir, white spruce, red spruce, black spruce, and along with non-host species including red maple (*Acer rubrum* L.), sugar maple (*Acer saccharum* Marsh.), American beech (*Fagus grandifolia*), yellow birch (*Betula alleghaniensis* Britton), paper birch (*Betula Papyrifera* Marsh.), trembling aspen (*Populus tremuloides* Michx.), and white cedar (*Thuja occidentalis*). The study area selected is mostly a part of the Canadian Acadian Forest Region (Rowe, 1972).

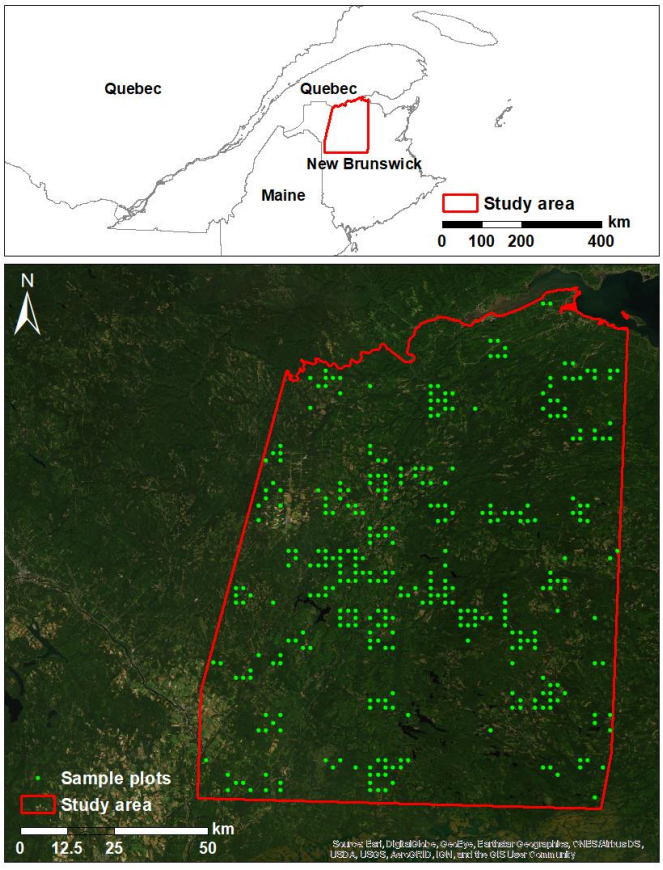


Figure 3.1. The extent of the study area and location of sample plots in NB, Canada.

3.2.2 Field Data

The field data were collected by the New Brunswick Department of Natural Resources and Energy Development (NBRED) under a continuous land inventory program. The inventory was conducted between 2016-2018 with a sample plot size of 400 m^2 and GPS accuracy of 10 cm (NBRED, 2015). We used a total of 349 samples with most plots being balsam fir or spruce dominant. Plots were classified into species classes based on percent basal area of species. A threshold of >70% combined basal area of the tree species present in a plot with mixed species was used to assign it to one of five classes: balsam fir (BF; 53 plots, 15%), balsam fir and broadleaved (BF-BL; 83 plots, 24%), BF and Spruce (BF-SP; 64 plots, 18%), broadleaved species (BL; 114 plots, 33%), and spruce (SP; 35 plots, 10%). The class BF-SP had more white and red spruce than black spruce, while SP had more black spruce. These classifications were based on prior research on species vulnerability to SBW (Hennigar et al., 2008) and observed distributions of the host species within the available data.

3.2.3 Remote Sensing Data Collection and Pre-processing

A total of 12 Sentinel-2A and -2B Level 1C Top-Of-Atmosphere (L1C-TOA) reflectance products (two scenes per date) from early June to early November were collected to represent the year-round phenology of the trees (https://earthexplorer.usgs.gov/). All leaf-on imagery was collected from June-September 2018. As cloud-free imagery was not available for fall 2018, two images from the years 2017 and 2019 were used instead (Table 3.1). Sentinel-2 L1C-TOA products were converted to L2A surface reflectance products at 20 m spatial resolution using the Sentinel Application Platform (SNAP) application with the Sen2Cor plugin before further processing.

Table 3.1. Remote sensing data used for this study.

Sensor	Imagery date	Resolution	Tile numbers/Absolute orbit number
Multispectral instrument (Sentinel-2)	June 6, 2018	Resampled to 20 m	T19UFP and T19TFN
	July 19, 2018		
	September 7, 2018		
	September 27, 2018		
	October 9, 2019		
	November 8, 2017		
C-band SAR (Sentinel-1)	May 11, 2018	Resampled to 20 m	021863
	June 4, 2018		022213
	July 22, 2018		022913
	September 20, 2018		023788
	November 7, 2018		024488

We obtained five tiles of Sentinel-1 Level 1 ground range detected (GRD) products from https://asf.alaska.edu/ for May to November 2018 (Table 3.1). SAR images for two polarizations (Vertical-Horizontal (VH) and Vertical-Vertical (VV)) were acquired using the Interferometric Wide (IW) Swath mode with a swath width of 250 km and 20 m × 5 m spatial resolution. Preprocessing of Sentinel-1 data was done using the SNAP tool in several steps: i) orbit file application to SAR images; ii) radiometric calibration; iii) speckle filter application; iv) terrain correction; and v) conversion of unit-less backscatter coefficient to decibel (Filipponi, 2019). All SAR images were resampled to 20 m.

3.2.4 Classification Algorithm and Variables

We used the Random Forest (RF) algorithm (Breiman, 2001) to develop a host species classification model because of promising results from previous research (Immitzer et al., 2016; Puletti et al., 2017; Persson et al., 2018; Grabska et al., 2019). RF is a non-parametric machine learning algorithm (an ensemble of decision trees) which makes the classification decision based on the votes from the trees grown using bootstrap samples of the data (Breiman, 2001). We used

the Caret package (Williams et al., 2015) in R to build, run and evaluate the RF model, with ten-fold cross-validation to assess model validity. The method splits the whole dataset into 10 parts, and each time nine parts are used for training the model and the rest for testing. Finally, the model averages the accuracy for all the iterations, which is the overall accuracy (OA) of the model.

RF was run with three datasets: i) Sentinel-1 and site variables; ii) Sentinel-2 and site variables; and iii) Sentinel-1, Sentinel-2 and site variables. For the predictor variables using Sentinel-2 satellite data, we used a combination of nine spectral bands (b2: Blue (490 nm), b3: Green (560 nm), b4: Red (665 nm), b5: Red-Edge (705nm), b6: Red-Edge (740 nm), b7: Red-Edge (783 nm), b8a: NIR (865 nm), b11: SWIR (1610 nm), b12: SWIR (2190 nm) and 20 SVIs including red-edge SVIs (Gitelson et al. 2002; Rahimzadeh-Bajgiran et al., 2012; Majasalmi and Rautiainen 2016, Bhattarai et al., 2020) and near infrared and shortwave infrared indices (Rahimzadeh-Bajgiran et al. 2018). Table 3.2 summarizes the Sentinel-2 based SVIs, for six different dates, resulting in a total of 174 optical variables. For Sentinel-1 SAR data, two variables (VV and VH polarizations) from five dates (a total of 10 SAR variables) were used.

Seven site variables including elevation, slope, aspect, soil types, Topographic Wetness Index (TWI), projected climate site index for year 2030 and improved Biomass Growth Index (iBGI; a potential site productivity measure) shown in Table 3.3, were used together with the satellite data. All site variables in raster format were resampled to 20 m spatial resolution.

We used the VSURF (Genuer et al., 2015) R-package to select the best variables from the available set of variables using the RF algorithm. Tree species of plots were grouped into five classes: BF, BF-BL, BF-SP, BL, and SP. The final model achieving the highest accuracy was used to create a host-species composition map at 20 m resolution.

Table 3.2. Sentinel-2 based SVIs used for the classification of SBW host species.

Acronym (SVIs)*	Formula	Reference
WDRVI	((0.01×b7)-b4)/((0.01×b7+b4)+ ((1-0.01) / (1+0.01)))	(Majasalmi and Rautiainen, 2016)
S2REP	705+35× (((b7+b4)/2-b5)/(b6-b5))	(Guyot and Baret, 1988)
NDVI	(b8a-b4)/(b8a+b4)	(Rouse et al., 1974)
NDVI45	(b5 - b4)/(b5 + b4)	(Delegido et al., 2011)
NDVI65	(b6 - b5)/(b6 + b5)	(Gitelson and Merzlyak, 1994)
MTCI	(b6-b5)/(b5-b4)	(Dash and Curran, 2007)
MSR	((b7/b4)-1)/√((b7/b4)+1)	(Chen, 1996)
IRECI	(b7-b4) × (b6/b5)	(Clevers et al., 2000)
GNDVI	(b8a-b3)/(b8a+b3)	(Gitelson et al., 1996)
EVI8	2.5× (b8a-b4)/(1+b8a+6×b4-7.5×b2)	(Huete et al., 2002)
EVI7	2.5× (b7-b4)/(1+b7+6×b4-7.5×b2)	(Majasalmi and Rautiainen, 2016)
ClRE	(b7/b5)-1	(Gitelson et al., 2003)
NDII11	(b8a-b11)/(b8a+b11)	(Hardisky et al., 1983)
NDII12	(b8a-b12)/(b8a+b12)	(Key et al., 2002)
ARI1	(1/b3)-(1/b5)	(Gitelson et al., 2001)
ARI2	(b8a/b3)-(b8a/b5)	
CRI1	(1/b2)-(1/b3)	(Gitelson et al., 2002)
CRI2	(1/b2)-(1/b5)	
MCARI	1-((0.2) × (b5-b3)/(b5-b4))	(Daughtry et al., 2000)
Red-Edge NDVI	(b8a-b6)/(b8a+b6)	(Gitelson and Merzlyak, 1994)

*: NIR: Near Infrared, SWIR: Shortwave Infrared, WDRVI: Wide Dynamic Range Vegetation Index, S2REP: Sentinel-2 Red Edge Position, NDVI: Normalized Difference Vegetation Index, MTCI: MERIS Terrestrial Chlorophyll Index, MSR: Modified Simple Ratio, IRECI: Inverted Red Edge Chlorophyll Index, GNDVI: Green NDVI, EVI: Enhanced Vegetation Index, ClRE: Chlorophyll Red Edge, NDII: Normalized Difference Infrared Index, ARI: Anthocyanin Reflectance Index, CRI: Carotenoid Reflectance Index, MCARI: Modified Chlorophyll Absorption in Reflectance Index

3.2.5 Balancing the Size of the Classes

Real-world datasets are prone to non-uniform class distributions, which tend to distort the performance of a classifier (Kurczab et al., 2014; Chaudhary et al., 2016) therefore balancing the sample size within classes significantly increases the accuracy of the model developed. Although the class-size balancing is often performed manually in the field of environmental and natural

sciences before training the model (He et al., 2019), there are several automated methods for balancing sample sizes. We balanced sample sizes of the imbalanced classes in our data using upSample and downSample functions available in R. The R-package SMOTE (Chawla et al., 2002) was used for this purpose which creates synthetic points to oversample the minority class while it randomly removes the observations from the majority class as per the arguments provided. Although this algorithm was created for binary classification, it can also be applied for multiple classes (Agrawal et al., 2015).

Table 3.3. Site variables used for the study.

Variables	Resolution	Reference/data provider
Elevation	30m resampled to 20 m	SRTM 30m void filled data
Slope	30m resampled to 20 m	SRTM 30m void filled data
Aspect	30m resampled to 20 m	SRTM 30m void filled data
Soil types	Vector	Hennigar et al. (2017)
Topographic wetness index (TWI)	30 m resampled to 20 m	Hennigar et al. (2017)
Projected climate site index (2030)	1 km resampled to 20 m	Jiang et al. (2015)
improved Biomass Growth Index (iBGI)	20 m	Rahimzadeh-Bajgiran et al. (2020)

*SRTM: Shuttle Radar Topography Mission

3.3 Results

3.3.1 Importance of Variables for Classification

Among the two SAR polarizations, VH had higher importance than VV polarization in the model. Among the 10 variables used from five different dates, the top Sentinel-1 variables were from May 11, 2018 (VH) and November 7, 2018 (VH). Figure 3.2a presents the rankings of the 10 most important variables for the model with Sentinel-1 and site variables. The addition of site variables did not yield marked differences in this model, except for elevation, TWI, and

iBGI. When evaluated with only four variables from the best dates, namely May 11 and November 7, 2018 (VH_11_7, VV_11_7, VH_05_11 and VV_05_11) with elevation, the model produced a similar accuracy as the one containing 10 variables. Given the similar importance of specific dates (Figure 3.2), only four SAR variables from May 11 and November 7, 2018, were used in the final model (Table 3.4).

Using Sentinel-2 and site variables, the 10 most influential variables included three variables from June 6, 2018 (b8a, b12, and IRECI), two variables from July 19, 2018 (ARI2 and b3), and the remaining five variables from October 9, 2019 (ClRE, MCARI, MSR, ARI1, and NDVI65) (Figure 3.2b). While all top SVIs were among red-edge SVIs, no traditional SVIs such as NDVI or EVI8 were found to be among top variables. Variables from other dates (September 7, 2018, September 27, 2018, and November 11, 2017) did not provide significant contributions to the final species classification model developed. Site variables were not among the top 10 variables and did not contribute to the performance of the model.

Combining Sentinel-1, Sentinel-2 and site variables, the 10 most influential variables were those identified in the Sentinel-2 model. None of the variables from Sentinel-1 SAR images, or site variables were among the top 10 important variables of the final model (Figure 3.2c).

Sentinel-1 mean backscatter intensity values for the five species classification classes are summarized in Figure 3.3. The spectral overlaps among the classes were more prominent in the May image than in the November image, and the backscatter separability was better in the VH amplitude band than in the VV band. As expected, the backscatter signature of the BL class was distinct from the other classes, especially using the VH band. Also, the backscatter values for the

BL class were higher in leaf-off conditions (November image) than during leaf-on conditions (May image) for both bands.

The Sentinel-2 mean spectral signatures for the five species classes over the three best dates (26 June, 19 July, and 9 October, 2018) were calculated and are presented in Figure 3.4. The reflectance values of the BL and BF-BL classes were higher than that of the pure coniferous classes mostly in the red-edge, NIR and SWIR regions, indicating good separability and potential for host species classification.

The mean spectral signatures also indicated that the BL class had a conspicuously different signature than that of other classes, while there were spectral overlaps between the BF, BF-SP, and SP classes. Reflectance of the BL class decreased towards the end of the fall season, while the other classes were similar in reflectance over time, as expected.

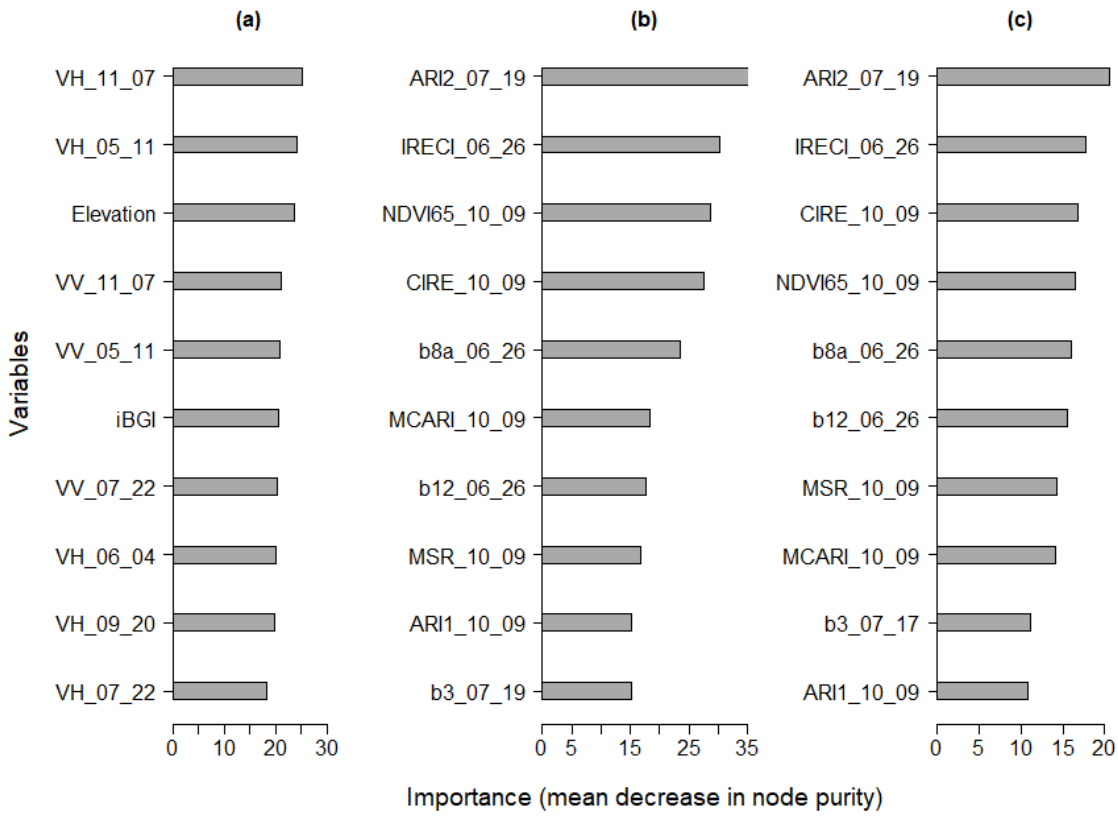


Figure 3.2. Importance ranking of the model variables from the Variable Selection Using Random Forest (VSURF) algorithm for the three different formulations based on the mean decrease in node purity: with a) Sentinel-1 and site variables, b) Sentinel-2 and site variables, and c) Sentinel-1, Sentinel-2 and site variables combined. Variable abbreviations follow the rule, name of the variable_acquisition month_day (see Table 3.2 and Table 3.3 for the further information).

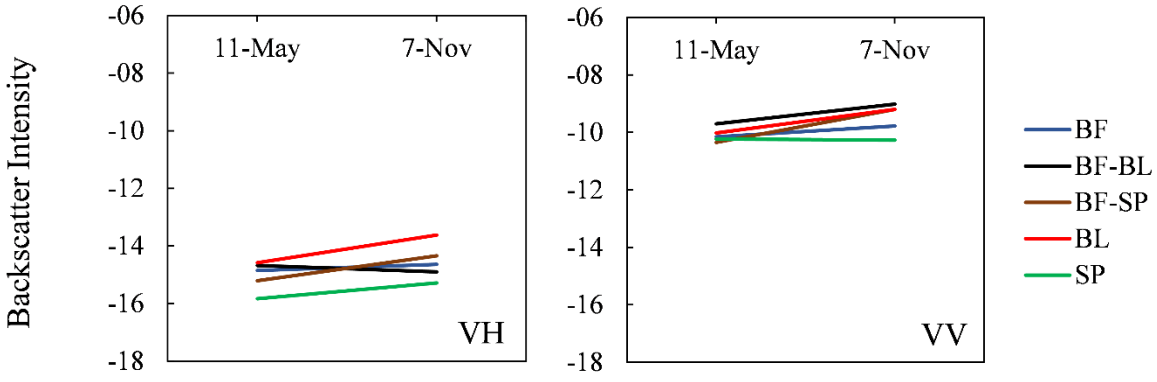


Figure 3.3. Sentinel-1 mean backscatter intensity values for VV and VH bands of the five classification classes over two extreme dates (leaf-on and leaf-off) for 2018. The tree species

classes were: i) balsam fir (BF); ii) balsam fir and broadleaved species (BF-BL); iii) balsam fir and spruce species mixed (BF-SP); iv) all broadleaved species (BL); and v) spruce species (SP).

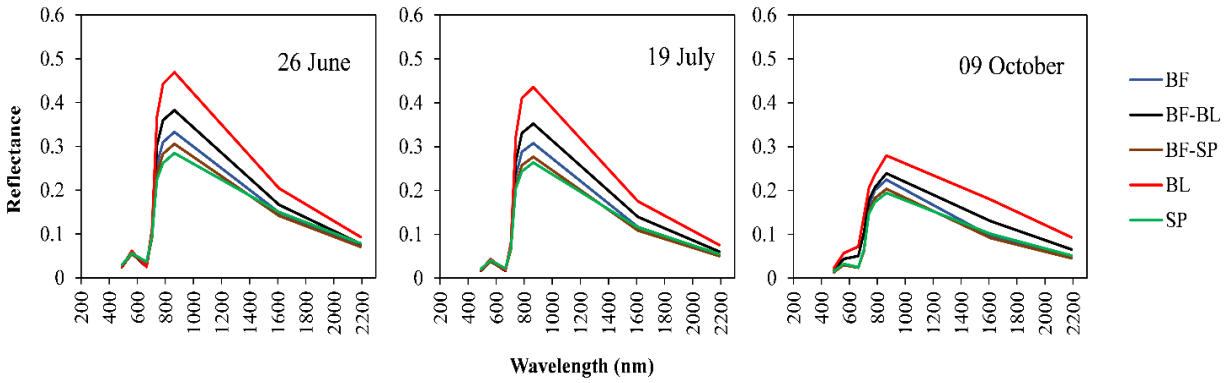


Figure 3.4. Sentinel-2 mean spectral signatures of five host tree species classes over three dates in 2018 (26 June and 19 July) and 2019 (9 October). The classes were: i) balsam fir (BF); ii) balsam fir and broadleaved species (BF-BL); iii) balsam fir and spruce species mixed (BF-SP); iv) all broadleaved species (BL); and v) spruce species (SP).

3.3.2 Classification Results and Model Validation

The overall accuracy (OA) and kappa coefficient (*K*) of the model built using Sentinel-1 images alone were 54% and 0.47, respectively, and including elevation increased the OA to 57.5% (Figure 3.2a). The addition of iBGI variable could not improve the model further. Using only the four variables from the best dates (May 11 and November 7, 2018) (VH_11_7, VV_11_7, VH_05_11 and VV_05_11) and elevation, the model still produced a similar accuracy (OA: 56% and *K*: 0.45). Therefore, only those variables were used in the next steps of the analysis (Table 3.4). Analyzing the confusion matrix (Table 3.4), the class-wise misclassification was more apparent among the host species (BF, BF-SP, SP) than between the host and non-host species. The performance of the model for the BL class was also low, indicating that Sentinel-1 variables were not very good at differentiating broadleaved species from balsam fir and spruce.

The OA and *K* obtained from the model based on Sentinel-2 data (Figure 3.2b) were 71% and 0.64 (Table 3.4). The class-wise confusion was prominent between the BF and SP classes, and other classes where BF is mixed (BF-BL and BF-SP). Also the BF-SP class was confused

with the BF and SP classes. In addition, confusion was also evident between the BL and BF-BL classes, however, the confusion between conifers and broadleaved species was minimal.

The combination of Sentinel-1, Sentinel-2 and site variables did not add much explanatory power to the model built with the Sentinel-2 variables alone (Figure 3.2b vs 3.2c). The inclusion of all 10 Sentinel-1 SAR variables with the top 10 Sentinel-2 variables (Figure 3.2a and 3.2c) improved the model performance by approximately 1% (OA: 72.2% and *K*: 0.646; Table 3.4). Reducing the number of SAR variables to only the two best variables from May and November images (VH-05-11 and VH-11-07) and elevation resulted in an OA of 72.3% and a *K* of 0.65 (Table 3.4).

Table 3.4. Confusion matrix and accuracy estimation (%) of the species classification models. PA is the producer's accuracy, and UA is the user's accuracy. Bold values are % correct classes.

	Species	BF	BF-BL	BF-SP	BL	SP	PA (%)	UA (%)
Sentinel-1 and elevation	BF	**56**	8	11	8	9	56	56
	BF-BL	12	**60**	11	12	15	60	58
	BF-SP	6	5	**45**	10	6	45	57
	BL	10	17	18	**58**	12	58	53
	SP	16	10	15	12	**58**	58	55
Overall accuracy: 56% and kappa coefficient: 0.45								
Sentinel-2	BF	**61**	6	18	0	6	61	63
	BF-BL	11	**76**	3	15	5	76	72
	BF-SP	17	2	**58**	0	13	58	61
	BL	2	14	0	**81**	1	81	83
	SP	9	2	21	4	**75**	75	73
Overall accuracy: 71.3% and kappa coefficient: 0.64								
Sentinel-1, Sentinel-2 and elevation	BF	**71**	6	12	3	4	71	71
	BF-BL	8	**66**	5	13	2	66	65
	BF-SP	11	6	**61**	0	14	61	63
	BL	0	16	0	**80**	0	80	88
	SP	10	6	22	4	**80**	80	70
Overall accuracy: 72.3% and kappa coefficient: 0.65								

The evaluation of performances of the three approaches illustrates that the model based on only Sentinel-2 optical bands and associated SVIs was the best for the SBW host species

classification, based on its simplicity and similar accuracy with the combined Sentinel-1 and Sentinel-2 model. The model with Sentinel-1 SAR and elevation performed considerably poorer than that with Sentinel-2 variables.

The resulting host species distribution maps are shown at two scales in Figure 3.5. Figure 3.5a was produced using the four best variables from Sentinel-1 SAR and elevation whereas Figure 3.5b shows the map produced using 10 best optical variables from Sentinel-2. Based on the map produced from Sentinel-2 model, of the total area of 810,500 ha, 11% was BF, 36% was BF-BL, 17% was BF-SP, 15% was BL and 20% was SP. Therefore most of the study area (over 80%) was susceptible to SBW defoliation from the view point of foliage availability.

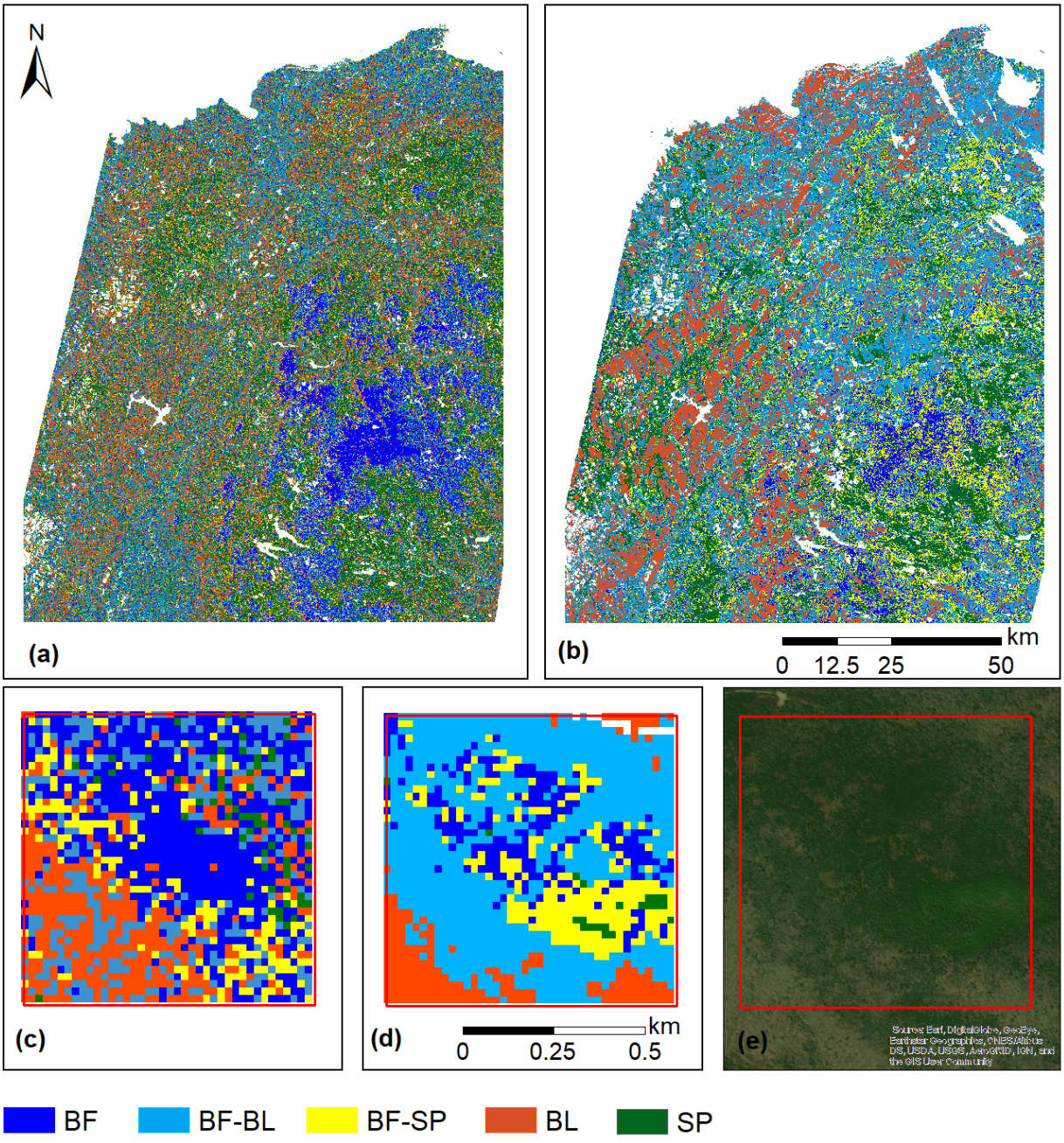


Figure 3.5. Host species distribution map at 20 m resolution using the RF model, (a) Sentinel-1 SAR images plus elevation, and (b) Sentinel-2 image variables. Maps (c), (d), and (e) are a subset of the study area extracted from map (a), (b), and Google Earth respectively (central

coordinates of the subsets are 47.15°N, and 66.66°W). White pixels are the clouds, water bodies, built-up areas and recent harvests.

3.4 Discussion

The individual application, as well as the integration of multispectral Sentinel-2 and Sentinel-1 SAR data with site variables, are novel approaches to construct a model for mapping SBW host tree species. Multispectral satellite sensors such as Landsat have been applied in various research works for mapping individual tree species or species groups. A combined application of satellite imagery and ancillary data is believed to result in more reliable outputs (Wolter et al., 1995; Wolter and Townsend, 2011). Our results showed only a marginal improvement in the performance of the classification model built with the integration of Sentinel-1, Sentinel- 2 and site variables compared to the performance of the model based on only Sentinel-2 variables (OA: 72.3% and *K*: 0.65 vs. OA: 71.3%, and *K*: 0.64). Sano et al. (2005) reported good performance of combined use of Japanese Earth Resources Satellite-1 (JERS-1) SAR and Landsat TM images for the discrimination of savannah vegetation ranging from grasslands to forests; however, they used L-Band SAR data. Moreover, their focus was on the discrimination of land cover not species composition within each land cover. Other similar studies also suggested the benefit of integrating SAR with optical multispectral images for vegetation classifications (Li et al., 2012; Erinjery et al., 2018; Schmidt et al., 2018; Yu et al., 2018). Although improvements were marginal in our study, the addition of Sentinel-1 variables did not deteriorate our model like it did in the study conducted by Rajah et al. (2019) for the detection and mapping of invasive species.

The inability of Sentinel-1 SAR variables to significantly improve the model with Sentinel-2 optical variables can also be attributed to the C-band wavelength, which cannot discriminate the subtle phenological changes in vegetation (due to lower canopy penetration

capability). The weakness with C-band could be improved using X-band (shorter wavelength SAR with higher canopy penetration capability) as suggested by Rajah et al. (2019), as the backscatter magnitude of SAR is dependent on its band frequency (Patel et al., 2006; Srivastava et al., 2009). In addition, the presence of heterogeneous species and their vertical structure affects the backscatter of SAR signals (Duguay et al., 2015), which is the case in our study as well. Li et al. (2012) also noted the incapability of C-band SAR images for the finer classification of vegetation in their attempt to classify different land cover types in northern Brazil.

The performance of the host species classification model using Sentinel-1 SAR variables alone was relatively satisfactory, but poorer than that of the model with Sentinel-2 variables. Among the two polarizations of Sentinel-1 SAR images, VH was better than VV in terms of its contribution to the model performance. The C-band SAR, especially at dual-polarization (VH in our case), is scattered by the foliage and branches of the upper canopy and has low crown penetration capacity, which makes it effective for coniferous and deciduous tree differentiation (Rignot et al., 1994). Rignot et al. (1994) used C, L, and P-band SAR (L and P bands are longer wavelength bands than C-band) to map parts of the Alaskan boreal forest and also concluded that HV (VH in our case) was the most useful polarization at all frequencies.

The inclusion of site variables did not improve the Sentinel-2 model, except for the TWI, and iBGI which marginally increased model performance. The species class accuracies from the Sentinel-2 model were 61% for BF, 76% for BF-BL, 58% for BF-SP, 81% for BL, and 75% for SP. High confusion was evident between the SP and BF species, while BL and SP species were better separable from other classes. Despite the confusion among the host species, our model is very efficient in separating non-host species from host species. Optical properties of the leaves

(interaction of electromagnetic waves and the leaf surface) for different tree species are unique and mostly dependent upon the photosynthetic as well as non-photosynthetic pigments available in the leaves, leaf water content and structure (external structure plus the orientation of mesophyll cells), leaf area index, leaf angle distribution and other factors (Ustin et al., 2009). The chemical and structural changes in BL tree canopy are significantly different than those of coniferous tree species over a year, which explains the accurate differentiation between them. However, both BF and SP species have needles and cone-like crown shapes creating difficulty in their separation. The crown shape and density of black spruce (cylindrical and open crown) is slightly different than that of white spruce (conical and dense), but the crown characteristics of white spruce resembles that of balsam fir (Honer, 1971).

The good performance of Sentinel-2 time series for the classification of SBW host species is in agreement with the prior studies utilizing Landsat time series (Wolter et al. 2008). However, Wolter et al. (2008) classified the host species into only two classes, BF and white spruce with an overall R^2 of 0.64, and 0.88, respectively. In contrast, our study classified the SBW host species into four categories, pure BF and SP (black spruce dominant), but also BF-SP (BF mixed with white and red spruce and BF-BL classes. The species classes we used are more consistent with SBW vulnerability and the available ground-based field data as there were rather limited pure stands.

Using multi-temporal images for modelling, we concluded that the late spring image (May or early June images) is an essential variable for the SBW host species discrimination. Similar results were found by Persson et al. (2018) and Grabska et al. (2019) for mapping forest species. Wolter and Townsend (2011) mapped the relative basal area of 12 forest species in northern Minnesota using multi-sensor data fusion, including commercial and non-commercial

optical (Landsat, and SPOT-5) and SAR (C-band Radarsat-1 and Phased Array type L-band (PALSAR)) satellite data and suggested the use of combined SAR (C-band Radarsat-1 and L-band PALSAR) and optical (Landsat and SPOT) data to achieve the highest accuracies (R^2=0.78, 0.80, and 0.93 for BF, white spruce, and black spruce, respectively).

The most important spectral variables for our model were anthocyanin indices (ARI1 and ARI2), red-edge indices (MSR, NDVI65, IRECI, ClRE and MCARI), NIR (b8a) and SWIR (b12) bands. SWIR, NIR, and red-edge bands and related indices have previously been used to predict leaf area index, canopy chlorophyll content and vegetation structure (Zhao et al., 2007; Majasalmi and Rautiainen 2016) successfully but not for tree species classification. Anthocyanin content can be taken as an indicator of leaf senescence in many plant species (Gitelson et al., 2001). The anthocyanin based indices' attaining high importance in our model can be attributed to the presence of both deciduous (leaf pigmentation during fall) and evergreen tree species in our study area. NIR, SWIR, and red-edge Sentinel-2 bands and SVIs were important in discriminating bramble (*Rubus cuneifolius*) (an invasive species) from surrounding vegetation using Sentinel-2 and Sentinel-1 SAR imagery in South Africa (Rajah et al., 2019). The importance of SWIR bands in SBW host mapping was also reported by Wolter et al. (2008), and the importance of red-edge and SWIR bands in tree species classification was noted by Immitzer et al. (2016) and Persson et al. (2018). Based on our research, red-edge SVIs are more effective than single spectral bands (including red-edge ones) and traditional SVIs for tree species classification.

Given the limited number of the studies on the SBW host species mapping using satellite imagery, the current study utilizing the latest and high-resolution open-source satellite data adds a useful contribution to facilitate future SBW monitoring. Our product can be used in

combination with LiDAR-based forest inventory data available for the study area as the latter do not provide accurate species identification information. The reliability of the model and the final host-composition map relies on the use of dense time-series images, red-edge SVIs (very sensitive to vegetation changes), the accuracy of the ground truth data, and a robust classification algorithm (RF). We used accurate ground dataset (~10 cm GPS accuracy), Sentinel-2 satellite time-series and 20 SVIs to produce a SBW host composition map at a resolution of 20 m. Further, this study also provided an alternative model derived from Sentinel-1 C-band SAR variables and elevation to produce a map of SBW host species with decent accuracy.

3.5 Conclusion

We compared models to classify SBW host species in NB, Canada, using Sentinel-1 and Sentinel-2 (alone and in combination), and seven site variables (elevation, slope, aspect, soil types, TWI, climate site index, and iBGI). The model derived from Sentinel-2 multi-temporal images using single spectral bands and SVIs outperformed other combinations. As anticipated, both Sentinel-1 and Sentinel-2 images acquired during the late spring and fall seasons were the key elements of our prediction models. The incorporation of Sentinel-1 SAR time-series with Sentinel-2 images did not markedly improve on the Sentinel-2 model.

Among the variables used in our study, the green, NIR and SWIR spectral bands and SVIs in particular red-edge SVIs were superior to other Sentinel-2 variables for tree species classification. Using Sentinel-1 SAR data, VH polarization was better for tree species discrimination than VV. The fine resolution and improved sensor properties of the new Sentinel-1 and Sentinel-2 satellites provide good capability for ground feature mapping and vegetation discrimination. Our results showed that Sentinel-2 multispectral data were better for tree species classification than Sentinel-1 SAR data.

CHAPTER 4

INTEGRATION, CONCLUSIONS AND RECOMMENDATIONS

In general, this study mapped spruce budworm (SBW) defoliation to provide some tools to help mitigate the probable loss of the commercially important spruce-fir forest. In this chapter, we incorporated the SBW host distribution map we produced from Sentinel-2 data in Chapter 3 with a SBW defoliation detection map derived from the best model (single-year RF) to see how the distribution of SBW defoliation is associated with the host availability. The previous chapters provided us with an opportunity to integrate the findings from them to compare the activity of SBW with host availability where study areas overlap keeping the fact in mind that SBW defoliation is also influenced by several other factors like SBW population density, host phenology, foliage characteristics, physical environment, etc. (Hennigar et al., 2008). Moreover the insect population depends upon the outbreak spread and insecticide treatments in recent years (MacLean et al., 2019). Further, we summarize our study and provide some recommendations in this section.

4.1 Integration of SBW Defoliation and Host Distribution Mapping

In Chapter 2, we compared the capabilities of available spectral bands and associated SVIs from Sentinel-2 satellite (20 SVIs and nine spectral bands) to detect and classify SBW defoliation. We selected five best variables from 29 available variables as remote sensing (RS) variables. From the site variables under consideration, only elevation contributed to improving the model significantly among others. We then combined five best RS variables obtained (EVI7, MCARI, IRECI, NDII11 and MSR for both defoliation detection and classification) and one site variable (elevation) to build the best model for defoliation detection and classification. The best

RS variables obtained to construct the model emphasize the importance of red-edge SVIs in identifying subtle changes in tree canopy structure and pigment content of the leaves due to SBW defoliation. The model further underlines the strong dependency of SBW defoliation on elevation (increased the classification model performance by 10%). Among the two machine learning algorithms (random forest (RF), and support vector machine (SVM)), and defoliation detection approaches (single-year, and multi-year) utilized for building detection and classification models in Chapter 2, single-year RF classification model performed better (at least more than 8%) than rest of the models. Consequently, it was considered for the production of defoliation detection and severity classification map. As the final product of Chapter 2, we produced a SBW defoliation severity map (Figure 4.1b) which demonstrates the concentration of SBW defoliation on the northern and central parts of New Brunswick (NB) for the year of 2018. The defoliation result obtained from Chapter 2 is the current-year defoliation and a function of outbreak spread from Quebec and heavy insecticide treatment under the early intervention strategy (EIS) program. Observing the pattern and severity of the defoliation, it is apparent that the severity of the attack is comparatively lower towards the southern part indicating that the defoliation has not yet progressed to these areas.

Likewise, in Chapter 3, we examined the performance of optical multispectral Sentinel-2 and Sentinel-1 SAR time series plus several site variables in discriminating the individual SBW host species among themselves and from the non-host species. We obtained similar results to those found to Chapter 2 indicating the superiority of SVIs compared to single spectral bands in building the model to map SBW host species distribution. The findings from Chapter 3 also revealed the influential performance of Sentinel-2 based time-series and red-edge SVIs in differentiating minor dissimilarities and pigment content on the canopies of different tree

species. The evaluation of different models built for species classification using Sentinel-1 SAR, Sentinel-2 optical multispectral and available site variables demonstrated the superior performance of the model (from the practical point of view) built with Sentinel-2 based variables alone. To compare with SBW defoliation map (Figure 4.1b), we presented a SBW host species distribution map derived from our best model for the same area in Figure 4.1a. The host species distribution map shows that balsam fir is widespread over the study area either in pure form or in combination with other species. The pure spruce is distributed more towards the south while broadleaved species are more common in the northern and southwestern part. Pure balsam fir is also seen more towards the southern part of the study area. Overall, the host species are available all over the study area in the form of pure stands or mixed stands.

Chapter 2 not only devised a model to detect and classify SBW defoliation but also provided us with an understanding of the pattern and severity of defoliation from a geographical perspective. The defoliation is evident mostly in the northern and central part of the study area. On the other hand, Chapter 3 assisted us to associate the severity and pattern of defoliation with the abundance of the host species in the same area. Evaluating the final products from Chapter 2 and Chapter 3, defoliation severity map and host species distribution map respectively (Figure 4.1), it is apparent that the SBW defoliation is following the distribution pattern of the host species. Among the host species, balsam fir is much more susceptible to SBW defoliation as compared to spruce species (Hennigar et al., 2008) which is evident in Figure 4.1 as well. However, the southern part of the study area, despite the presence of host species seems undefoliated. The low defoliation in southern part might be a result of the EIS program in NB from 2014 to 2018. MacLean et al. (2018) show the highest area of forest in northern New Brunswick (199,696 ha) being treated with insecticides or pheromone as compared to previous

years which might have controlled the SBW population in the hotspots and prevented them from spreading. Therefore, the combination of these two chapters helps us to verify the degree of impact laid by the distribution and abundance of primary host species on the severity of the SBW defoliation.

Comparing the two maps in Figure 4.1, we can notice the defoliation occurring in the areas with higher amount of fir trees supporting the findings from Hennigar et al. (2008) and Sainte-Marie et al. (2015) about host priority of SBW (balsam fir being the first priority host of SBW). Evaluating the defoliation severity across the four host classes (BF, BF-BL, BF-SP and SP) it is observed that most of the areas are not defoliated yet. Of the total balsam fir area, the percentage coverage of “No defoliation”, “Light” and “Moderate” were 70%, 5% and 25%, respectively. Similarly the percentage coverage for “No defoliation”, “Light” and “Moderate” for “balsam fir and broadleaved”, “balsam fir and spruce”, and “spruce” were (55.5%, 7.5% and 37%), (69%, 5.5% and 25.5%), and (87%, 4%, and 7%), respectively. Figure 4.2 demonstrates the SBW defoliation taking place in various classes of host species on varying severity levels.

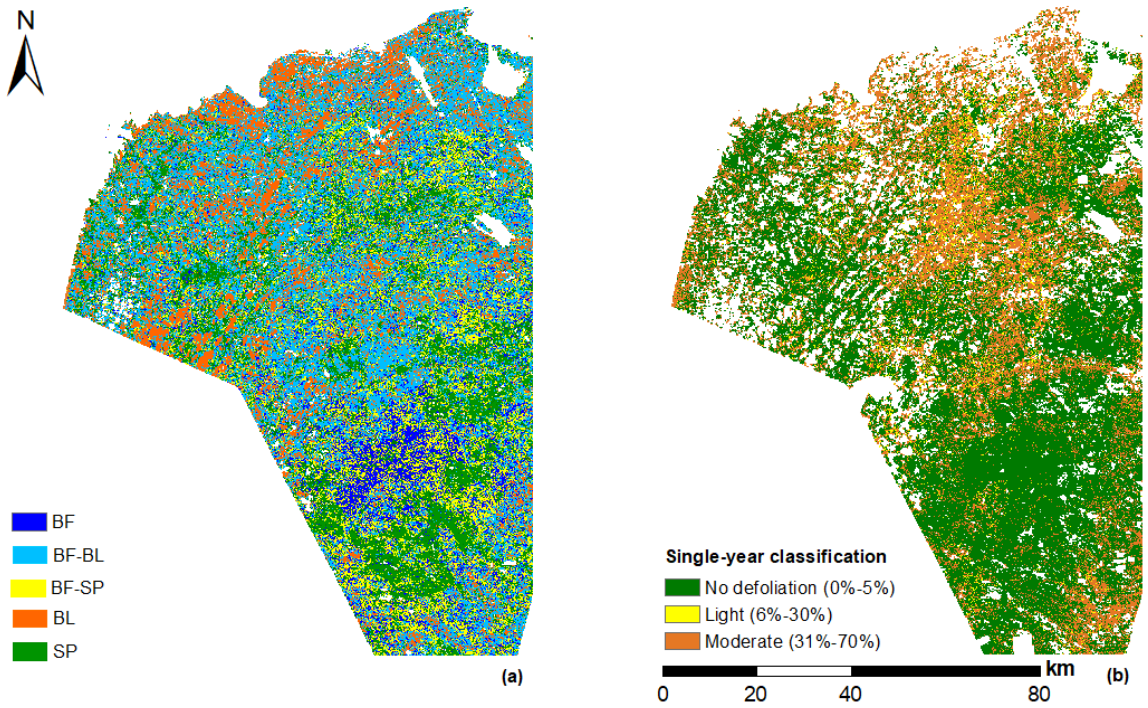


Figure 4.1. (a). The SBW host species composition map using RF model from Sentinel-2 variables at 20 m. (b). SBW defoliation severity classification map using single-year RF model at 20 m.

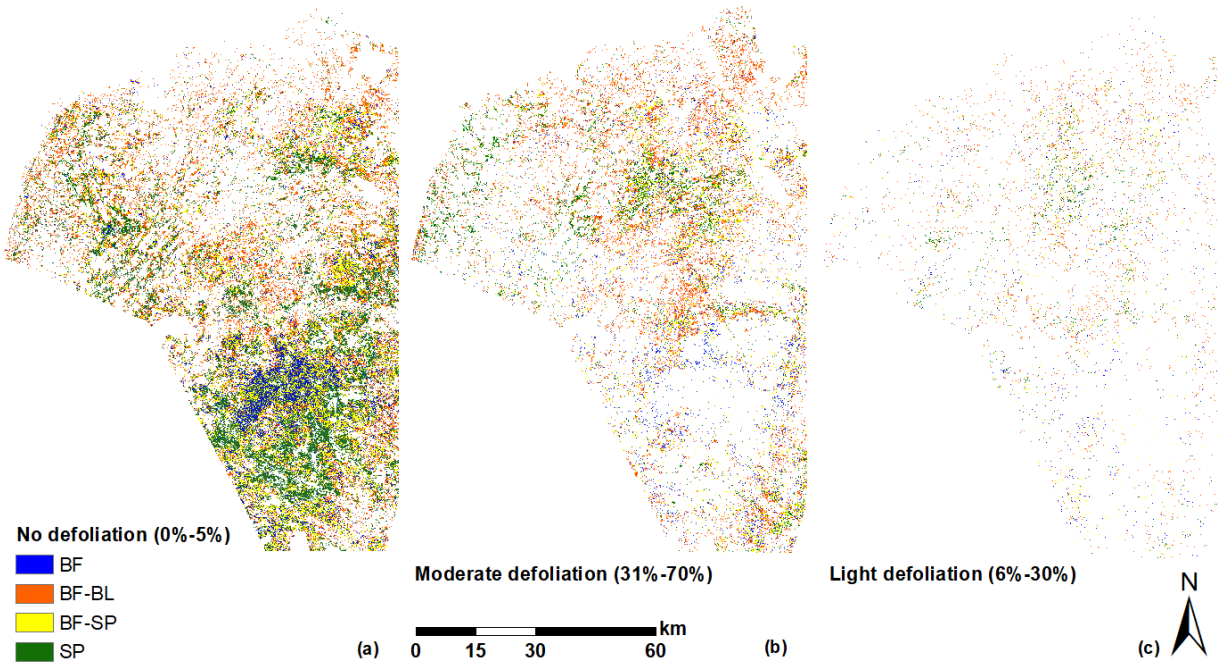


Figure 4.2. Defoliation by host species classes across various severity classes, (a). No defoliation, (b). Light defoliation, and (c). Moderate defoliation.

4.2 Conclusion and Recommendations

SBW defoliation is a recurring phenomenon in Northeastern forests of the United States and Canada, and its prevention is crucial from both economic and ecological points of view. The conventional way of sketching the defoliated area from an aircraft seems comparatively uneconomic and tedious considering the current rise of cost-effective high-resolution satellite sensors. Although there is considerable potential for using remote sensing to address forest health issues, it is still in its evolving stage of development and requires further research; however, the outcomes of ongoing research are still promising.

The accuracy of remote sensing techniques not only relies upon the sensor capabilities but also is influenced by the extent to which the ground truth data are associated with the sensor information (the accuracy of field data). Therefore, the use of accurate ground sampling techniques coherent with the particular sensor specifications is strongly recommended prior to

the implementation of remote sensing techniques for ecological studies. Particularly, for SBW defoliation, current-year defoliation classification is crucial for the early intervention strategies from a practical viewpoint. Therefore the ground survey, as well as the satellite data collection, should be timed within the two-week window during which current-year SBW defoliation can be detected. Using optical multispectral satellite data cloud contamination has always been a challenge, but the high temporal resolution of the twin Sentinel-2 satellites will provide better quality data for SBW defoliation detection.

The detection and classification of light SBW defoliation is still challenging even with the use of high-resolution satellite images like Sentinel-2 at 20 m spatial resolution. In these circumstances there exists a high probability of confusion between light and non-defoliated pixels when defoliation is scattered in small patches. The combined use of suggested models to produce annual SBW defoliation and recent host species maps could be of great value for assisting the monitoring of SBW defoliation along with traditional monitoring methods. The proposed technique operates with low cost and human resource requirement as compared to using conventional methods like aerial sketch mapping, which demand a considerable amount of human resources and money.

Given the heterogeneous composition of the forest in our study area, the distinct phenology of different tree species is an essential phenomenon to capture for differentiating those species using remote sensing techniques. Hence, we used and suggest a time series (both leaf off and leaf on images) of satellite images for species differentiation over using a single-date image. Further, red-edge bands and red-edge SVIs are influential in picking the subtle changes in tree canopy structure and pigment content and play a vital role both in the detection of SBW defoliation and discrimination of SBW host species. The new red-edge indices, which we can

derive from the Sentinel-2 data are much more dominant and sensitive towards the subtle changes in vegetation pigments and canopy structure than conventional SVIs like NDVI. Therefore we suggest the utilization of valuable open source Sentinel-2 data for similar studies on forest health and tree species mapping. Finally, it is commendable to acquire adequate information on the distribution of SBW host species to reinforce the monitoring of SBW defoliation because they are strictly host-specific and attack the spruce and balsam fir species only.

The two sources of satellite data (Sentinel-1 and Sentinel-2) used in our study proved useful for SBW host species classification and could be used individually to model the distribution of tree species. However, their combined use did not augment the accuracy of the model considerably as expected and is not recommended for similar future endeavors at least using the same satellite data preprocessing techniques approached in this study. Similarly, for SBW defoliation severity classification, Sentinel-2 imagery and derived SVIs were invaluable resources for our work and should be explored more in future research.

REFERENCES

Adelabu, S., Mutanga, O., Adam, E. E., and Cho, M. A. (2013). Exploiting machine learning algorithms for tree species classification in a semiarid woodland using RapidEye image. Journal of Applied Remote Sensing, 7(1), 073480.

Adelabu, S., Mutanga, O., and Cho, M. A. (2012). A review of remote sensing of insect defoliation and its implications for the detection and mapping of *Imbrasia belina* defoliation of Mopane Woodland. The African Journal of Plant Science and Biotechnology, 6, 1-13.

Agrawal, A., Viktor, H. L., and Paquet, E. (2015). SCUT: Multi-class imbalanced data classification using SMOTE and cluster-based undersampling. Paper presented at the 2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K). (Vol. 1, pp. 226-234). IEEE.

Barra, A., Solari, L., Béjar-Pizarro, M., Monserrat, O., Bianchini, S., Herrera, G., Crosetto, M., Sarro, R., González-Alonso, E., Mateos, R.M. and Ligüerzana, S. (2017). A methodology to detect and update active deformation areas based on sentinel-1 SAR images. Remote Sensing, 9(10), 1002.

Bennetts, R. E., White, G. C., Hawksworth, F. G., and Severs, S. E. (1996). The influence of dwarf mistletoe on bird communities in Colorado ponderosa pine forests. Ecological Applications, 6(3), 899-909.

Bhattarai, R., Rahimzadeh-Bajgiran, P., Weiskittel, A. and MacLean, D. (2020). Sentinel-2 based prediction of spruce budworm defoliation using red-edge spectral vegetation indices. Remote Sensing Letters, 11 (8), 777-786.

Blais, J. R. (1983). Trends in the frequency, extent, and severity of spruce budworm outbreaks in eastern Canada. Canadian Journal of Forest Research, 13(4), 539-547.

Bouchard, M., Kneeshaw, D., and Bergeron, Y. (2006). Forest dynamics after successive spruce budworm outbreaks in mixedwood forests. Ecology, 87(9), 2319-2329.

Boulanger, Y., Arseneault, D., Morin, H., Jardon, Y., Bertrand, P., and Dagneau, C. (2012). Dendrochronological reconstruction of spruce budworm (*Choristoneura fumiferana*) outbreaks in southern Quebec for the last 400 years. Canadian Journal of Forest Research, 42(7), 1264-1276.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

Brovkina, O., Cienciala, E., Zemek, F., Lukeš, P., Fabianek, T., and Russ, R. (2017). Composite indicator for monitoring of Norway spruce stand decline. European Journal of Remote Sensing, 50(1), 550-563.

Carleer, A., and Wolff, E. (2004). Exploitation of very high resolution satellite data for tree species identification. Photogrammetric Engineering and Remote Sensing, 70(1), 135-140.

Chaudhary, A., Kolhe, S., and Kamal, R. (2016). An improved random forest classifier for multi-class classification. Information Processing in Agriculture, 3(4), 215-222.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321-357.

Chen, C., Rahimzadeh-Bajgiran, P., and Weiskittel, A. (Under preparation). Assessing spatial and temporal dynamics of a spruce budworm outbreak across the complex forested landscape of Maine, USA.

Chen, C., Weiskittel, A., Bataineh, M. and MacLean, D.A. (2017a). Evaluating the influence of varying levels of spruce budworm defoliation on annualized individual tree growth and mortality in Maine, USA and New Brunswick, Canada. Forest Ecology and Management, 396, 184-194.

Chen, C., Weiskittel, A., Bataineh, M. and MacLean, D.A. (2017b). Even low levels of spruce budworm defoliation affect mortality and ingrowth but net growth is more driven by competition. Canadian Journal of Forest Research, 47(11), 1546-1556.

Chen, C., Weiskittel, A., Bataineh, M., and MacLean, D. A. (2018). Modelling variation and temporal dynamics of individual tree defoliation caused by spruce budworm in Maine, US and New Brunswick, Canada. Forestry: An International Journal of Forest Research, 92(1), 133-145.

Chen, J. M. (1996). Evaluation of vegetation indices and a modified simple ratio for boreal applications. Canadian Journal of Remote Sensing, 22(3), 229-242.

Ciesla, W. M. (2000). Remote sensing in forest health protection. FHTET Report No. 3. US Department of Agriculture, Forest Service, Forest Health Technology Enterprise Team, Remote Sensing Applications Center, 266.

Ciesla, W. M., Dull, C. W., and Acciavatti, R. E. (1989). Interpretation of SPOT-1 color composites for mapping defoliation of hardwood forests by gypsy moth. Photogrammetric Engineering and Remote Sensing, 55(10), 1465-1470.

Clevers, J., De Jong, S., Epema, G., Addink, E., Van Der Meer, F., and Skidmore, A. (2000). Meris and the Red-edge index. Paper presented at the Second EARSeL workshop on Imaging spectroscopy. EARSeL, Enschede.

Clevers, J. G., and Kooistra, L. (2011). Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5(2), 574-583.

Dale, V., Gardner, R., DeAngelis, D., Eagar, C., and Webb, J. (1991). Elevation-mediated effects of balsam woolly adelgid on southern Appalachian spruce–fir forests. Canadian Journal of Forest Research, 21(11), 1639-1648.

Dale, V. H., Joyce, L. A., McNulty, S., Neilson, R. P., Ayres, M. P., Flannigan, M. D., Hanson, P.J., Irland, L.C., Lugo, A.E., Peterson, C.J. and Simberloff, D. (2001). Climate change and forest disturbances: climate change can affect forests by altering the frequency, intensity, duration, and timing of fire, drought, introduced species, insect and pathogen outbreaks, hurricanes, windstorms, ice storms, or landslides. BioScience, 51(9), 723-734.

Damask, J. N. (2004). Polarization optics in telecommunications (Vol. 101): Springer Science and Business Media.

Dash, J., and Curran, P. (2007). Evaluation of the MERIS terrestrial chlorophyll index (MTCI). Advances in Space Research, 39(1), 100-104.

Daughtry, C., Walthall, C., Kim, M., De Colstoun, E. B., and McMurtrey III, J. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment, 74(2), 229-239.

De Beurs, K., and Townsend, P. (2008). Estimating the effect of gypsy moth defoliation using MODIS. Remote Sensing of Environment, 112(10), 3983-3990.

Delegido, J., Verrelst, J., Alonso, L., and Moreno, J. (2011). Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. Sensors, 11(7), 7063-7081.

Duguay, Y., Bernier, M., Lévesque, E., and Tremblay, B. (2015). Potential of C and X band SAR for shrub growth monitoring in sub-arctic environments. Remote Sensing, 7(7), 9410-9430.

Eklundh, L., Johansson, T., and Solberg, S. (2009). Mapping insect defoliation in Scots pine with MODIS time-series data. Remote Sensing of Environment, 113(7), 1566-1573.

Erdle, T.A and Ward, C. (2008). Management alternatives for New Brunswick's public forest. Report of the New Brunswick Task Force on Forest Diversity and Wood Supply. Fredericton. Retrieved from https://www2.gnb.ca/content/dam/gnb/Departments/nrrn/pdf/en/ForestsCrownLands/Erdle/ErdleReport-e.pdf

Erinjery, J. J., Singh, M., and Kent, R. (2018). Mapping and assessment of vegetation types in the tropical rainforests of the Western Ghats using multispectral Sentinel-2 and SAR Sentinel-1 satellite imagery. Remote Sensing of Environment, 216, 345-354.

Eschen, R., O'Hanlon, R., Santini, A., Vannini, A., Roques, A., Kirichenko, N., and Kenis, M. (2019). Safeguarding global plant health: the rise of sentinels. Journal of Pest Science, 92(1), 29-36.

European Space Agency (2016). Sentinel-1 User Guide. Retrieved from https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar

Fan, H. (2006). Satellite remote sensing of cumulative spruce budworm defoliation in Prince Albert National Park (Doctoral dissertation, University of Saskatchewan).

Filipponi, F. (2019). Sentinel-1 GRD Preprocessing Workflow. Paper presented at the Multidisciplinary Digital Publishing Institute Proceedings (Vol. 18, No.1, p.11).

Fleming, R. A., Candau, J.N., and McAlpine, R. S. (2002). Landscape-scale analysis of interactions between insect defoliation and forest fire in central Canada. Climatic Change, 55(1-2), 251-272.

Frampton, W. J., Dash, J., Watmough, G., and Milton, E. J. (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. ISPRS Journal of Photogrammetry and Remote Sensing, 82, 83-92.

Genuer, R., Poggi, J.M., and Tuleau-Malot, C. (2015). VSURF: an R package for variable selection using random forests. The R Journal, 7(2), 19-33.

Giri, C., Pengra, B., Long, J., and Loveland, T. R. (2013). Next generation of global land cover characterization, mapping, and monitoring. International Journal of Applied Earth Observation and Geoinformation, 25, 30-37.

Gitelson, A., and Merzlyak, M. N. (1994). Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. Journal of Photochemistry Photobiology B: Biology, 22(3), 247-252.

Gitelson, A. A., Kaufman, Y. J., and Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment, 58(3), 289-298.

Gitelson, A. A., Merzlyak, M. N., and Chivkunova, O. B. (2001). Optical properties and nondestructive estimation of anthocyanin content in plant leaves. Photochemistry and Photobiology, 74(1), 38-45.

Gitelson, A. A., Viña, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., and Leavitt, B. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. Geophysical Research Letters, 30(5), 52-1-52-4.

Gitelson, A. A., Zur, Y., Chivkunova, O. B., and Merzlyak, M. N. (2002). Assessing Carotenoid Content in Plant Leaves with Reflectance Spectroscopy. Photochemistry Photobiology, 75(3), 272-281.

Grabska, E., Hostert, P., Pflugmacher, D., and Ostapowicz, K. (2019). Forest stand species mapping using the Sentinel-2 time series. Remote Sensing, 11(10), 1197.

Guyot, G., and Baret, F. (1988). Utilisation de la haute resolution spectrale pour suivre l'etat des couverts vegetaux. Paper presented at the Spectral Signatures of Objects in Remote Sensing, 287, 279.

Hall, R., Castilla, G., White, J., Cooke, B., and Skakun, R. (2016). Remote sensing of forest pest damage: a review and lessons learned from a Canadian perspective. The Canadian Entomologist, 148(S1), S296-S356.

Hall, R. J., Skakun, R. S., and Arsenault, E. J. (2006). Remotely sensed data in the mapping of insect defoliation. Understanding forest disturbance and spatial pattern: Remote sensing and GIS approaches, Chap. 4, 85-111.

Hardisky, M., Klemas, V., and Smart, M. (1983). The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of *Spartina alterniflora* canopies. Photogrametric Engineering and Remote Sensing, 49(1), 77-83.

Hawryło, P., Bednarz, B., Wężyk, P., and Szostak, M. (2018). Estimating defoliation of Scots pine stands using machine learning methods and vegetation indices of Sentinel-2. European Journal of Remote Sensing, 51(1), 194-204.

He, Y., Chen, G., Potter, C., and Meentemeyer, R. K. (2019). Integrating multi-sensor remote sensing and species distribution modeling to map the spread of emerging forest disease and tree mortality. Remote Sensing of Environment, 231, 111238.

Hennigar, C. R., MacLean, D. A., Erdle, T. A., and Wagner, R. (2013). Potential spruce budworm impacts and mitigation opportunities in Maine. Report submitted to Cooperative Forest Research Unit (CFRU), University of Maine, Orono.

Hennigar, C., Weiskittel, A., Allen, H. L., and MacLean, D. A. (2017). Development and evaluation of a biomass increment based index for site productivity. Canadian Journal of Forest Research, 47(3), 400-410.

Hennigar, C. R., MacLean, D. A., Quiring, D. T., and Kershaw Jr, J. A. (2008). Differences in spruce budworm defoliation among balsam fir and white, red, and black spruce. Forest Science, 54(2), 158-166.

Honer, T.G. (1971). Crown shape in open-and forest-grown balsam fir and black spruce. Canadian Journal of Forest Research, 1(4), 203-207.

Huang, Z. (2015). Remote sensing of spruce budworm defoliation in Quebec, Canada using eo-1 hyperion data (MScE thesis, University of New Brunswick).

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment, 83(1-2), 195-213.

Immitzer, M., Atzberger, C., and Koukal, T. (2012). Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. Remote Sensing, 4(9), 2661-2693.

Immitzer, M., Vuolo, F., and Atzberger, C. (2016). First experience with Sentinel-2 data for crop and tree species classifications in central Europe. Remote Sensing, 8(3), 166.

Jepsen, J., Hagen, S., Høgda, K., Ims, R., Karlsen, S., Tømmervik, H., and Yoccoz, N. (2009). Monitoring the spatio-temporal dynamics of geometrid moth outbreaks in birch forest using MODIS-NDVI data. Remote Sensing of Environment, 113(9), 1939.

Jiang, H., Radtke, P.J., Weiskittel, A.R., Coulston, J.W. and Guertin, P.J. (2015). Climate-and soil-based models of site productivity in eastern US tree species. Canadian Journal of Forest Research, 45(3), 325-342.

Johns, R.C., Bowden, J.J., Carleton, D.R., Cooke, B.J., Edwards, S., Emilson, E.J., James, P., Kneeshaw, D., MacLean, D.A., Martel, V. and Moise, E.R. (2019). A Conceptual Framework for the Spruce Budworm Early Intervention Strategy: Can Outbreaks be stopped? Forests, 10(10), 910.

Jones, H.G. and Vaughan, R.A., 2010. Remote sensing of vegetation: principles, techniques, and applications. Oxford University Press.

Karasiak, N., Sheeren, D., Fauvel, M., Willm, J., Dejoux, J.F., and Monteil, C. (2017). Mapping tree species of forests in southwest France using sentinel-2 image time series. Paper presented at the 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp) (pp. 1-4).IEEE.

Kettela, E. G. (1982). Results of aerial surveys for current spruce budworm defoliation in New Brunswick and a review of the methodology. Environ. Can., Can. For. Serv., Maritimes For. Res. Centre, Fredericton, NB. Tech. Note, (66), 8.

Kettela, E. G. (1983). A cartographic history of spruce budworm defoliation from 1967 to 1981 in Eastern North America (Vol. 14).

Key, C., Benson, N., Ohlen, D., Howard, S., and Zhu, Z. (2002). The normalized burn ratio and relationships to burn severity: ecology, remote sensing and implementation. Paper presented at the Ninth biennial remote sensing applications conference, Apr 8–12, San Diego, CA.

Kuenzer, C., Ottinger, M., Wegmann, M., Guo, H., Wang, C., Zhang, J., Dech, S. and Wikelski, M. (2014). Earth observation satellite sensors for biodiversity monitoring: potentials and bottlenecks. International Journal of Remote Sensing, 35(18), 6599-6647.

Kurczab, R., Smusz, S., and Bojarski, A. J. (2014). The influence of negative training set size on machine learning-based virtual screening. Journal of Cheminformatics, 6(1), 32.

Kurosu, T., Uratsuka, S., Maeno, H., and Kozu, T. (1999). Texture statistics for classification of land use with multitemporal JERS-1 SAR single-look imagery. IEEE Transactions on Geoscience and Remote Sensing, 37(1), 227-235.

Kurosu, T., Yokoyama, S., Fujita, M., and Chiba, K. (2001). Land use classification with textural analysis and the aggregation technique using multi-temporal JERS-1 L-band SAR images. International Journal of Remote Sensing, 22(4), 595-613.

Leckie, D. G., Yuan, X., Ostaff, D. P., Piene, H., and MacLean, D. (1992). Analysis of high resolution multispectral MEIS imagery for spruce budworm damage assessment on a single tree basis. Remote Sensing of Environment, 40(2), 125-136.

Legaard, K., Simons-Legaard, E. and Weiskittel, A. (2020). Multi-Objective Support Vector Regression Reduces Systematic Error in Moderate Resolution Maps of Tree Species Abundance. Remote Sensing, 12(11), 1739.

Li, G., Lu, D., Moran, E., Dutra, L., and Batistella, M. (2012). A comparative analysis of ALOS PALSAR L-band and RADARSAT-2 C-band data for land-cover classification in a tropical moist region. ISPRS Journal of Photogrammetry and Remote Sensing, 70, 26-38.

Li, X., and Yeh, A. (2004). Multitemporal SAR images for monitoring cultivation systems using case-based reasoning. Remote Sensing of Environment, 90(4), 524-534.

Lin, C.F. (2004). Training algorithms for fuzzy support vector machines with noisy data. Pattern Recognition Letters, 25(14), 1647-1656.

MacLean, D. A. (1980). Vulnerability of fir-spruce stands during uncontrolled spruce budworm outbreaks: a review and discussion. The Forestry Chronicle, 56(5), 213-221.

MacLean, D. A. (1984). Effects of spruce budworm outbreaks on the productivity and stability of balsam fir forests. The Forestry Chronicle, 60(5), 273-279.

MacLean, D. A., Amirault, P., Amos-Binks, L., Carleton, D., Hennigar, C., Johns, R., and Régnière, J. (2019). Positive results of an early intervention strategy to suppress a spruce budworm outbreak after five years of trials. Forests, 10(5), 448.

MacLean, D. A., and MacKinnon, W. E. (1996). Accuracy of aerial sketch-mapping estimates of spruce budworm defoliation in New Brunswick. Canadian Journal of Forest Research, 26(12), 2099-2108.

MacLean, D. A., and Piene, H. (1995). Spatial and temporal patterns of balsam fir mortality in spaced and unspaced stands caused by spruce budworm defoliation. Canadian Journal of Forest Research, 25(6), 902-911.

Main, R., Mathieu, R., Kleynhans, W., Wessels, K., Naidoo, L., and Asner, G. P. (2016). Hyper-temporal C-band SAR for baseline woody structural assessments in deciduous savannas. Remote Sensing, 8(8), 661.

Majasalmi, T., and Rautiainen, M. (2016). The potential of Sentinel-2 data for estimating biophysical variables in a boreal forest: a simulation study. Remote Sensing Letters, 7(5), 427-436.

Miller, C. (1977). The feeding impact of spruce budworm on balsam fir. Canadian Journal of Forest Research, 7(1), 76-84.

Ministère des Forêts de la Faune et des Parcs (2019). Aires infestées par la tordeuse des bourgeons de l'épinette au Québec en 2019, Québec, Gouvernement du Québec, Direction de la protection des forêts, 32. https://mffp.gouv.qc.ca/publications/forets/fimaq/insectes/tordeuse/TBE_2019_P.pdf

Morin, H. (1998). Importance et évolution des épidémies de la tordeuse des bourgeons de l'épinette dans l'est du Canada: l'apport de la dendrochronologie. Géographie Physique et Quaternaire, 52(2), 237-244.

Morris, R., and Bishop, R. (1951). A method of rapid forest survey for mapping vulnerability to spruce budworm damage. The Forestry Chronicle, 27(2), 171-178.

Nelson, M. (2017). Evaluating Multitemporal Sentinel-2 Data for Forest Mapping Using Random Forest (Master's dissertation, Stockholm University, Stockholm, Sweden).

Pal, M. (2005). Random forest classifier for remote sensing classification. International Journal of Remote Sensing, 26(1), 217-222.

Patel, P., Srivastava, H. S., Panigrahy, S., and Parihar, J. S. (2006). Comparative evaluation of the sensitivity of multi-polarized multi-frequency SAR backscatter to plant density. International Journal of Remote Sensing, 27(2), 293-305.

Persson, M., Lindberg, E., and Reese, H. (2018). Tree species classification with multi-temporal Sentinel-2 data. Remote Sensing, 10(11), 1794.

Peterson, D., Resetar, S., Brower, J., and Diver, R. (1999). Forest monitoring and remote sensing: a survey accomplishments for the opportunities for the future. RAND Science and Technology Policy Institute, Washington DC, USA.

Puletti, N., Chianucci, F., and Castaldi, C. (2017). Use of Sentinel-2 for forest classification in Mediterranean environments. Annals of Silvicultural Research, 42(1), 32-38.

Rahimzadeh-Bajgiran, P., Hennigar, C., Weiskittel, A. and Lamb, S. (2020). Forest Potential Productivity Mapping by Linking Remote-Sensing-Derived Metrics to Site Variables. Remote Sensing, 12(12), 2056.

Rahimzadeh-Bajgiran, P., Munehiro, M., and Omasa, K. (2012a). Relationships between the photochemical reflectance index (PRI) and chlorophyll fluorescence parameters and plant pigment indices at different leaf growth stages. Photosynthesis Research, 113(1-3), 261-271.

Rahimzadeh-Bajgiran, P., Omasa, K., and Shimizu, Y. (2012b). Comparative evaluation of the Vegetation Dryness Index (VDI), the Temperature Vegetation Dryness Index (TVDI) and the improved TVDI (iTVDI) for water stress detection in semi-arid regions of Iran. ISPRS Journal of Photogrammetry and Remote Sensing, 68, 1-12.

Rahimzadeh-Bajgiran, P., Weiskittel, A., Kneeshaw, D., and MacLean, D. (2018). Detection of annual spruce budworm defoliation and severity classification using Landsat imagery. Forests, 9(6), 357.

Rajah, P., Odindi, J., Mutanga, O., and Kiala, Z. (2019). The utility of Sentinel-2 Vegetation Indices (VIs) and Sentinel-1 Synthetic Aperture Radar (SAR) for invasive alien species detection and mapping. Nature Conservation, 35, 41-46.

Ranson, K. J., and Sun, G. (1994). Northern forest classification using temporal multifrequency and multipolarimetric SAR images. Remote Sensing of Environment, 47(2), 142-153.

Rignot, E. J., Williams, C. L., Way, J., and Viereck, L. A. (1994). Mapping of forest types in Alaskan boreal forests using SAR imagery. IEEE Transactions on Geoscience and Remote Sensing, 32(5), 1051-1059.

Rose, A. H., and Lindquist, O. (1985). Insects of eastern spruces, fir and hemlock. Ottawa:Canadian Forestry Service Technical Report 23.

Rouse Jr, J. W., Haas, R., Schell, J., and Deering, D.W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. NASA Special Publication, 351,309.

Rowe, J. S. (1972). Forest regions of Canada, Publication No. 1300, Canadian Forestry Service, Department of the Environment, Ottawa, Ontario.

Royama, T.O. (1984). Population dynamics of the spruce budworm *Choristoneura fumiferana*. Ecological Monographs, 54(4), 429-462.

Sainte-Marie, G. B., Kneeshaw, D. D., MacLean, D. A., and Hennigar, C. R. (2015). Estimating forest vulnerability to the next spruce budworm outbreak: will past silvicultural efforts pay dividends? Canadian Journal of Forest Research, 45(3), 314-324.

Sano, E. E., Ferreira, L. G., and Huete, A. R. (2005). Synthetic aperture radar (L band) and optical vegetation indices for discriminating the Brazilian savanna physiognomies: A comparative analysis. Earth Interactions, 9(15), 1-15.

Schmidt, J., Fassnacht, F. E., Förster, M., and Schmidtlein, S. (2018). Synergetic use of Sentinel-1 and Sentinel-2 for assessments of heathland conservation status. Remote Sensing in Ecology, 4(3), 225-239.

Silva, C. R., Olthoff, A., de la Mata, J. A. D., and Alonso, A. P. (2013). Remote monitoring of forest insect defoliation. A review. Forest Systems, 22(3), 377-391.

Simmons, M., Davis D.,Griffiths L., and Muecke A. (1984). The natural history of Nova Scotia. Nova Scotia Department of Education and Nova Scotia Department of Lands and Forests. Volume 1 and 2.

Srivastava, H. S., Patel, P., Sharma, Y., and Navalgund, R. R. (2009). Multi-frequency and multi-polarized SAR response to thin vegetation and scattered trees. Current Science, 97(3), 425-429.

Statnikov, A., Wang, L., and Aliferis, C. F. (2008). A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. BMC Bioinformatics, 9(1), 319.

Su, Q., Needham, T.D. and MacLean, D.A. (1996). The influence of hardwood content on balsam fir defoliation by spruce budworm. Canadian Journal of Forest Research, 26(9), 1620-1628.

Thapa, B., Wolter, P.T., Sturtevant, B.R. and Townsend, P.A. (2020). Reconstructing past forest composition and abundance by using archived Landsat and national forest inventory data. International Journal of Remote Sensing, 41(10), 4022-4056.

Taylor, S. L., and MacLean, D. A. (2008). Validation of spruce budworm outbreak history developed from aerial sketch mapping of defoliation in New Brunswick. Northern Journal of Applied Forestry, 25(3), 139-145.

Tewkesbury, A. P., Comber, A. J., Tate, N. J., Lamb, A., and Fisher, P. F. (2015). A critical synthesis of remotely sensed optical image change detection techniques. Remote Sensing of Environment, 160, 1-14.

Tkacz, B., Moody, B., and Villa Castillo, J. (2007). Forest health status in North America. The Scientific World Journal, 7, 28-36.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M. and Traver, I.N. (2012). GMES Sentinel-1 mission. Remote Sensing of Environment, 120, 9-24.

Townsend, P. A., Singh, A., Foster, J. R., Rehberg, N. J., Kingdon, C. C., Eshleman, K. N., and Seagle, S. W. (2012). A general Landsat model to predict canopy defoliation in broadleaf deciduous forests. Remote Sensing of Environment, 119, 255-265.

Ustin, S.L., Gitelson, A.A., Jacquemoud, S., Schaepman, M., Asner, G.P., Gamon, J.A. and Zarco-Tejada, P. (2009). Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sensing of Environment, 113, S67-S77.

Van Beijma, S., Comber, A., and Lamb, A. (2014). Random forest classification of salt marsh vegetation habitats using quad-polarimetric airborne SAR, elevation and optical RS data. Remote Sensing of Environment, 149, 118-129.

Wagner, R. G., Bryant, J., Burgason, B., Doty, M., Roth, B. E., Strauch, P., . . . Denico, D. (2015). Coming spruce budworm outbreak: Initial risk assessment and preparation and response recommendations for Maine's forestry community. Cooperative Forestry Research Unit, University of Maine, Orono. 77p.

Waters, W., Heller, R., and Bean, J. (1958). Aerial appraisal of damage by the spruce budworm. Journal of Forestry, 56(4), 269-276.

Weier, J. and Herring, D., 2000. Measuring vegetation (NDVI and EVI). NASA Earth Observatory, 20. Retrieved from http://earthobservatory.nasa.gov/Features/MeasuringVegetation/

White, J. D., Kroh, G. C., and Pinder, J. E. (1995). Forest mapping at Lassen Volcanic National Park, California, using Landsat TM data and a geographical information system. Photogrammetric Engineering and Remote Sensing, 61(3), 299-305.

Williams, C. K., Engelhardt, A., Cooper, T., Mayer, Z., Ziem, A., Scrucca, L., Tang, Y., Candan, C., Hunt, T., Kuhn, and M. M. (2015). 'caret'. R Package. Available online: https://pbil.univ-lyon1.fr/CRAN/web/packages/caret/caret.pdf.

Wolter, P.T., Mladenoff, D.J., Host, G.E. and Crow, T.R. (1995). Improved forest classification in the Northern Lake States using multi-temporal Landsat imagery. Photogrammetric Engineering and Remote Sensing, 61(9), 1129-1143.

Wolter, P. T., and Townsend, P. A. (2011). Multi-sensor data fusion for estimating forest species composition and abundance in northern Minnesota. Remote Sensing of Environment, 115(2), 671-691.

Wolter, P. T., Townsend, P. A., Sturtevant, B. R., and Kingdon, C. C. (2008). Remote sensing of the distribution and abundance of host species for spruce budworm in Northern Minnesota and Ontario. Remote Sensing of Environment, 112(10), 3971-3982.

Yu, Y., Li, M., and Fu, Y. (2018). Forest type identification by random forest classification combined with SPOT and multitemporal SAR data. Journal of Forestry Research, 29(5), 1407-1414.

Zhang, B., MacLean, D.A., Johns, R.C. and Eveleigh, E.S. (2018). Effects of hardwood content on balsam fir defoliation during the building phase of a spruce budworm outbreak. Forests, 9(9), 530.

Zhao, D., Huang, L., Li, J., and Qi, J. (2007). A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. ISPRS Journal of Photogrammetry and Remote Sensing, 62(1), 25-33.

BIOGRAPHY OF THE AUTHOR

Rajeev Bhattarai was born in Bandipur, Tanahun, Nepal on 10th August 1994 and raised in the same place. He went to Notre Dame School (operated by School Sisters of Notre Dame) for his primary, secondary, and high school (graduated from the high school in 2012). Rajeev was admitted to Institute of Forestry, Pokhara Campus, Nepal for his undergraduate degree (2013- 2017). Other than academia, he enjoys sports and leadership. During his undergraduate, he, directly and indirectly, involved abundantly in leading various organizations within the campus as well as games and sports activities.

Rajeev worked in several rural areas of Nepal after his graduation as a field assistant for a couple of projects like Project 111 (joint research conducted by Chiba University of Japan and Institute of Forestry, Nepal), and Asia Network for Sustainable Agriculture and Bioresources. He enjoys mingling with the crude culture of rural people rather than just enjoying the developments in the cities.

Rajeev is a candidate for Masters of Science degree in Forest Resources from the University of Maine in August 2019.