Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Early detection of emerald ash borer infestation using multisourced data: a case study in the town of Oakville, Ontario, Canada

Kongwen Zhang
Baoxin Hu
Justin Robinson

SPIE.

Early detection of emerald ash borer infestation using multisourced data: a case study in the town of Oakville, Ontario, Canada

Kongwen Zhang,[a,b,*] Baoxin Hu,[a] and Justin Robinson[b]
[a]York University, Department of Earth and Space Science and Engineering, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada
[b]Selkirk Geospatial Research Centre, 301 Frank Beinder Way, Castlegar, British Columbia V1N 4L3, Canada

Abstract. The emerald ash borer (EAB) poses a significant economic and environmental threat to ash trees in southern Ontario, Canada, and the northern states of the USA. It is critical that effective technologies are urgently developed to detect, monitor, and control the spread of EAB. This paper presents a methodology using multisourced data to predict potential infestations of EAB in the town of Oakville, Ontario, Canada. The information combined in this study includes remotely sensed data, such as high spatial resolution aerial imagery, commercial ground and airborne hyperspectral data, and Google Earth imagery, in addition to nonremotely sensed data, such as archived paper maps and documents. This wide range of data provides extensive information that can be used for early detection of EAB, yet their effective employment and use remain a significant challenge. A prediction function was developed to estimate the EAB infestation states of individual ash trees using three major attributes: leaf chlorophyll content, tree crown spatial pattern, and prior knowledge. Comparison between these predicted values and a ground-based survey demonstrated an overall accuracy of 62.5%, with 22.5% omission and 18.5% commission errors.
[DOI: 10.1117/1.JRS.8.083602]

Keywords: emerald ash borer; early detection; hyperspectral; data analysis; remote sensing.

Paper 14076 received Feb. 3, 2014; revised manuscript received May 13, 2014; accepted for publication Jun. 2, 2014; published online Jul. 7, 2014.

1 Introduction

The emerald ash borer (*Agrilus planipennis*, EAB) is one of the most destructive insects to affect ash species of the genus Fraxinus.[1,2] Ash is one of the most popular landscape trees in North America, replacing elm trees in new residential and commercial developments due to its high tolerance to environmental stresses and its resistance to pests.[3] The invasion of EAB was first reported in 2002, in Michigan, USA. Despite substantial research and control efforts, the beetle has continued to spread to new areas, and today it is found in 14 states and has crossed the border to southern Ontario, Canada.[4]

The beetle has caused the death of millions of ash trees and billions of dollars in economic losses.[5] The ash trees attacked by EABs often show serious decline within 2 years of initial infestation, and typically die within 3 to 5 years.[6] As such, early detection is especially valued due to the aggressive nature of the EAB infestation.

Infestations are normally identified by tree owners and are confirmed by onsite tree branch sampling.[7] Recent EAB studies have focused on biochemical and biological symptoms and treatments, such as tree leaf decline, woodpecker presence, chemical injection treatments, and tree removal.[3] However, each of these techniques is inherently concerned with late stage occurrences of the infestation. In order to improve this, multiple government agencies and environmental groups are exploring the use of remote sensing techniques for infestation detection over

*Address all correspondence to: Kongwen Zhang, E-mail: conwen@yorku.ca

large areas, including city and provincial scales. Analysis of such data could offer workable results with improved cost and detectability efficiencies.

During EAB infestation, the larvae destroy the cambium layer under the bark responsible for transporting nutrients and water throughout the tree, therefore resulting in general stress.[7] Trees under general stress often experience changes in leaf pigment composition, particularly in leaf chlorophyll content. Thus, from a detection point of view, trees exhibiting general stress symptoms are much more likely to be infested.[8] As a result, spectral features that are sensitive to leaf chlorophyll content can be potentially used for detecting EAB infestations with remote sensing technologies.[9] Current remote sensing approaches are focusing on such spectral features, but their success has been limited, as described below.

In Pontius et al.,[9] a six-parameter function was proposed, and a 97% correlation was reported between selected vegetation indices and tree decline caused by EAB infestations. However, this correlation can only be confirmed in the dieback stages, and the spectral-based variation is not sufficiently significant during the early infestation stage to allow unique identification.

Souci et al.[10] presented an EAB detection project based in Milwaukee, Wisconsin, using high spatial resolution hyperspectral images and a pixel-based vegetation index (VI) approach to map out the invasive species. They also used light detection and ranging (LiDAR) to create a digital elevation model and applied geographic information system (GIS) buffer zone tools to highlight potential infested areas. However, no significant validation result was obtained, and this study is considered as a pioneer work in multisource data combination.

Hanou[11] used a VI to represent the correlation between infestation and tree crown pixels to map out ash infestation in Oakville, Ontario. It can be expressed as

$$\frac{(R_{\mathrm{SWIR}} + R_{\mathrm{Green}})^2}{R_{\mathrm{SWIR}} - R_{\mathrm{Green}}}, \tag{1}$$

where R_{SWIR} is the short wavelength infrared band and R_{Green} is the green band. However, this pixel-based VI approach can result in the misclassification of different tree species as ash, and struggles to identify tree canopy over different backgrounds, as well as being prone to spectral variations within individual trees.[11] As a result, it was not able to successfully map infested ash trees using this method.

Our goal in this study is to make this difficult task of early stage EAB infestation identification possible, providing a methodology that can deliver a reasonable detectability for EAB infestation over large spatial scales. After background searches and the literature reviews, we have determined that the "clues" associated with EAB infestation are widely varied and may be deeply buried in different data sources, and not just spectral data from hyperspectral imagery. As such, spectral and pixel-based analysis alone is not sufficient for early detection of EAB, as it is susceptible to considerable in-class spectral variation. It is necessary for us to broaden our range of information sources, which requires the precise identification of information correlated with EAB infestation and their effective retrieval from different data sources. This additional information from varied sources can reduce the ambiguity and uncertainties caused by in-class spectral variation and ultimately improve the accuracy of information retrieval.

First, we investigated spatial features as a potential information source. As suggested by Smiteley et al.,[12] there are a number of visual symptoms that can help to detect early stage infestation, such as the dieback of twigs, thinning of the crown, and growth of new branches on the lower trunk. These symptoms can often be reflected in structural changes of the tree crowns.[13,14] For a mature and healthy ash tree, an oval-shaped crown can be assumed. In the presence of EABs, an ash tree loses its leaves, starting from the top central branches, which leads to "holes" or "gaps" in the crown. In a high spatial resolution image, a healthy mature ash tree exhibits a smooth texture with "salt and pepper" patterning, which represents the gaps between leaves and branches. In contrast, certain patterns that derive from the presence of larger "holes" in the crown are expected in an EAB infested tree. Since spectral features do not typically consider the spatial patterning changes before and after infestation, this predictive information on tree morphology may help to determine the EAB infestation level.

We also recognize the natural spreading capacity of EAB. The mature beetle can migrate from one tree to its neighbors within a 25 km radius.[15,16] Therefore, the known infested

trees that have been previously documented are important contributors to the potential continuation of the infestation, which is a consideration that has been long neglected.

Further, following a comprehensive literature search and preliminary studies and testing of various data sources, we identify the leaf chlorophyll content as a potential indicator of infestation. Specifically, a significant drop compared to the surroundings, in addition to the tree crown degradation and distance from known infested trees may be considered key factors that can potentially lead to accurate estimation of the extent of infestation.

Despite recognition of these key factors, it remains a challenge to effectively assemble the diverse information derived from multisourced data in order to reach a meaningful, consistent, and accurate conclusion. As a solution to this, we propose an object-oriented approach rooted in these key factors, and devise a weighted linear prediction score function. With proper calibration, this new infestation scoring approach can be applied to the detection of early to medium stage EAB infestations.

Therefore, the overall outline of this study is as follows, also shown in detail in Table 1.

1. Comprehensive background search and the literature review.
2. Extensive data collection and processing.
3. Prior processing studies and testing.
4. Identification of the key factors for EAB infestation, which must be strongly correlated with infestation and must be extractable from our data.
5. Establish an operational workflow, which can effectively and robustly retrieve information from different sources and intelligently combine them to deliver a precise prediction.
6. Map out the health of individual ash trees within test areas using the proposed methodology.

2 Data Specification and Preprocessing

2.1 *Study Area*

The study area was in the town of Oakville, Ontario (Fig. 1). All ash trees within the township boundary were part of the study. However, due to data limitations, particularly related to ground

Table 1 Overall structure of this study.

	Spectral features	Spatial features	Prior knowledge
Background and the literature review	Pontius et al., (2008) [9]	Smiteley et al., (2008) [12]	BenDor et al., (2006) [15]
	Souci et al., (2009) [10] Hanou (2010) [11], etc.	Zhang and Hu (2012) [17]	Herms et al., (2009) [16]
Data sources	Hyperspectral data:	High spatial resolution data:	Community maps and paper documents
	1) Ground ASD measurements;	1) YUL Oakville imagery;	
	2) Airborne imagery	2) Google Earth imagery	
Prior processing studies and testing	1) Ground ASD spectral signature separability study 2) Contribution analysis for vegetation indices band selection	Spatial pattern visual interpretation	EAB natural spread capability
Key factors identified	Leaf chlorophyll content	Tree crown degradation	Distance from nearby known infested trees
Operational workflow	1) Three vegetation indices 2) Leaf chlorophyll content retrieved from model inversion	Longitudinal profiles	Distance constant
Map out the health of individual ash trees			

Fig. 1 Location of the town of Oakville, Ontario, Canada, and the ground survey results of individual tree distribution. Numbers in the inset are the IDs of trees with confirmed infestations.

data acquisition, we focused mainly on six sites within Oakville, which are indicated in the final infestation map. Oakville is the first town in Canada to use aerial-based imagery to compile and inventory public and private ash tree damage.[11] Oakville is located in the EAB infestation zone, and has multiple confirmed cases of EAB infestation. Therefore, this region has great interest in research into early detection of EAB for both environmental and economical reasons.

2.2 *Airborne and Ground Hyperspectral Data*

Two sets of commercial hyperspectral remote sensing data were used:

1. Airborne hyperspectral imagery with a spatial resolution of 1 m was collected using a ProSpectTIR-VS2 in July 2010. In total, 360 spectral bands were recorded, ranging from 390 to 2400 nm with a band spacing of 5 nm. The raw imagery was converted into radiance by removing dark current, and was then corrected for atmospheric disturbances using ATCOR4 software. Precise georeferencing and geometric correction was carried out using information provided by the onboard internal navigation system and global positioning system.[11]

2. Ground-level hyperspectral data were also collected. The leaves from ash trees at different stages of infestation were measured using an analytical spectral device (ASD) (ASD Inc., Boulder, Colorado, USA). Leaves were recorded as healthy, low, medium, and highly infested. This is a fundamental prior processing study, used for separability assessment, which must be conducted before any further meaningful research can be begun. If no significant differences are observed between ground level spectra of trees at different stages of infestation, then the likelihood of canopy-level spectra exhibiting sufficient differences for classification is very low.

2.3 *High Spatial Resolution Aerial Imagery*

High spatial resolution aerial imagery was provided by the York University library. The data were collected during the winter of 2008, having 1 m resolution, and are orthorectified with blue, red, and green in three spectral bands. However, the absence of leaves at the time of image acquisition limited the use of these data in this study. We, therefore, only use the images to validate the locations of trees and to confirm tree species identification.

2.4 *Google Earth Aerial Imagery*

The employment of Google Earth (GE) aerial imagery in scientific studies poses a dilemma. GE imagery is not quality assured with any known or documented geometric or radiometric corrections. However, the data are free and easy to access, and are potentially very powerful, containing, for example, multiple years of data from the same location. In the Oakville area, GE imagery has an average spatial resolution of 13–15 cm over our test sites. Before making use of the GE imagery, we made precise calculations of scale using known objects, such as rooftops. We also validated the geometric locations of trees, and the differences in the locations of our sampled trees. From this, we found that the GE imagery and the orthorectified images were within 0.5 m of the tree centers. We did not validate or calibrate the radiometric values of the GE imagery, as we were only concerned with the relative changes in the pixels.

2.5 *Ground Truthing*

We used two sources of ground-based information to determine the likelihood of infestation from nearby infested or healthy trees, and to support ground validation. First, most infestations were confirmed using a previously documented branch sampling approach.[7] Two branches, approximately 50 cm long and no less than 6 cm in diameter, of each targeted ash tree were obtained, their bark was peeled off, and the underlying cambium inspected for evidence of EAB infestation, such as serpentine galleries, larvae, and young beetles (Fig. 2).

Second, nine community maps showing street and park tree layouts were supplied by the town of Oakville. In these maps, the infestation states of individual trees were recorded as "healthy," "low," "medium," and "high" infestation states, as indicated by ground-based survey. We manually transferred these locations and condition information into our imagery data.

3 Operational Workflow and Detail Procedures

We have identified three key factors that can potentially guide an estimation of the health of ash trees, supported by previous work and preliminary studies. These are the leaf chlorophyll content, tree crown degradation, and distance from nearby trees of known health state. An operational workflow must be established, with the objectives of: (1) effectively retrieving the targeted information from its associated data source; and (2) intelligently assembling all information to provide a final estimation.

Despite the convenience of pixel-based methods,[9–11] we proposed an object-oriented approach, which ultimately forms a score-based prediction function. In this study, individual tree crowns were considered as the basic objects, instead of isolated pixels, and their spectral and spatial properties were analyzed. Using an object-oriented approach in this way is essential for the retrieval of distance and crown degradation information, while also improving the computing efficiency in the model inversion.

Fig. 2 The branch sampling approach for EAB detection. A larva is found in this sample. Copyrighted at AMEC Inc., used with permission.

Our approach is comprised of four major steps: segmentation and identification of individual ash trees; information retrieval; information conversion; and final evaluation through score calculation. This operational workflow is graphically depicted in Fig. 3. As shown in this figure, the airborne hyperspectral and the GE aerial imagery data underwent segmentation to outline tree boundaries. Then, the vegetation indices, leaf chlorophyll content derived from model inversion, and distance constants, originally from maps but transferred to imagery tree locations, were derived from the hyperspectral data, whereas the spatial patterns of crown degradation were obtained from the GE imagery. All "information" from this step was converted into normalized numerical values between 0 and 5, and assembled into a score function, which provides a prediction score that can be used to determine the health state of individual trees.

3.1 *Segmentation and Identification of Individual Ash Trees*

To ensure the accuracy of the locations and boundaries of individual ash tree crowns, multiscale automatic segmentation[18] results from both the hyperspectral and high spatial resolution images were manually checked and corrected. In the hyperspectral image, the pixels within each tree crown were identified and classified into the same crown. Using ground surveys and the community maps provided by Oakville, all ash trees used in the following training and validation studies were checked and confirmed.

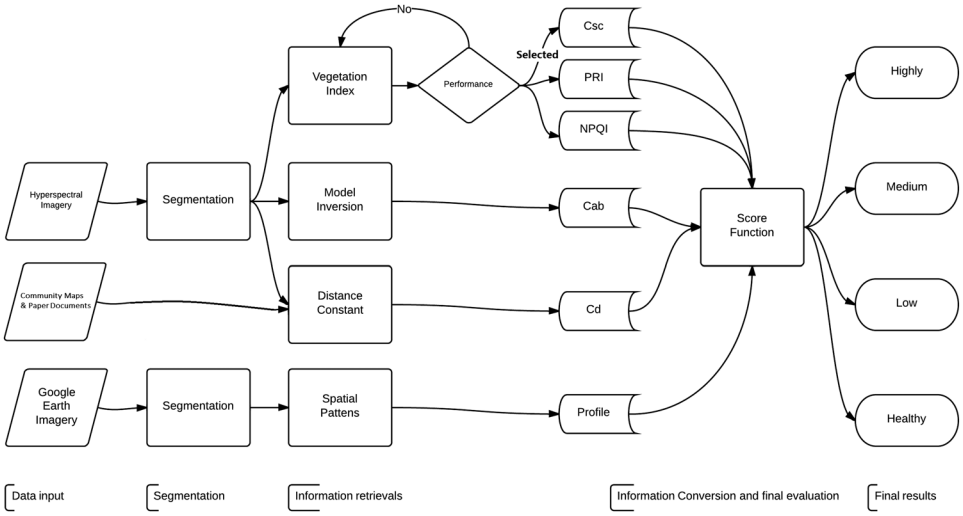

Fig. 3 Graphical depiction of the operational workflow for this study, showing four major steps retrieving four types of information and six different parameters from three different data sources to derive the final health estimation.

3.2 *Information Retrieval*

3.2.1 *Overview*

As shown in Fig. 3, four types of information, making up six indices or parameters, need to be retrieved from the varied data sources, such as three vegetation indices: the leaf chlorophyll content; the distance constant; and the longitudinal profile. The following sections will detail the individual methods used to extract each of these parameters.

3.2.2 *Calculation of the selected vegetation indices related to the health state of tree crowns from spectral features*

Attempts have been previously made by many researchers to identify the health state of tree crowns from airborne and satellite images, through the design of a VI sensitive to biochemical properties of leaves (e.g., Refs. 9, 19, and 20). In these studies, four major categories of biochemical parameters have been considered: leaf chlorophyll content,[9,19,21] carotenoids,[20,22] water content,[23] and anthocyanins.[22] After a comprehensive review of these properties and their associated vegetation indices, a suite of indices was selected (Table 2) as candidates for detailed comparison and future selection. The existing indices were typically designed for specific sites and sensors,[9,19,22,24] and it was found that using the same formula with a small change in the wavelength used in the index could result in a noticeable difference in the correlations between a VI and its respective vegetation parameter.[19]

We applied a model-based site-independent band selection analysis called contribution analysis, theoretically equivalent to a sensitive study, to optimize the selection of vegetation indices. The contribution analysis calculates the contribution index (CI), which provides detailed, quantitative measurements of the correlations between each observation band and the target parameters, solely based on the model that we use.[31] For this study, we calculated the CI for each spectral band based on the PROSPECT leaf model[32] due to the prevailing deciduous leaf type. PROSPECT is well validated and is known for successfully simulating deciduous trees situations. The model simulates leaf reflectance and transmittance at 400–2500 nm and takes four key input parameters, such as the leaf chlorophyll content, the water content, dry matter, and a parameter describing the internal structure of the leaf, denoted as N. We calculated CI using PROSPECT, which covers all the spectral bands for VIs listed in Table 2. CI is the most important criterion for VI band selection, and the bands that have higher CIs are preferred over the lower scoring ones. In addition to this, we also conducted an actual performance evaluation. We used half of the known infested trees as a training data set, described in detail in the training section below, and the remainders were used as validators to evaluate the performance of the various VIs. From these analyses, the three top performing indices were chosen for the final infestation evaluation based on the following criteria: (1) higher CI in those bands involved in the VI; (2) higher separability in different health states; and (3) lower in-class variation, which is quantified by the coefficient of variation (CV). CV measures the dispersion of the measurement, which is the extent of variability in relation to the mean of a population. For VI selection, this means that VI should have minimum class variation (e.g., low standard deviation) in each state. Essentially, the greater the separation shown between different stages, particularly the early and median stage, the better the candidate.

After determining the final three optimum VIs (PRI, CSc, and NPQI; the selection criteria are described in detail in later sections), we used them to calculate VI values from the hyperspectral spectra extracted from the airborne hyperspectral imagery. We used the average spectra of the sunlit portion of the tree crown only, but included any dark spots within that area.

3.2.3 *Retrieving leaf chlorophyll content based on physical model inversion from spectral features*

Physical model inversion is a sophisticated approach that can obtain accurate biophysical parameter estimations.[31] The observed and model-simulated spectra are compared using a predefined merit function to determine the best-estimated model parameters.

Table 2 Vegetation indices selected for investigation in this study.

	Vegetation index formula	Parameter sensitive to	References
Chlorophyll and stress			
Csc	$R605/R760$	Chlorophyll a	[24]
GI	$R554/R677$	LAI and total chlorophyll content	[25]
Vogb	FD715/FD705	Total chlorophyll content	[26]
NPQI	$(R415 - R435)/(R415 + R435)$	Chlorophyll degradation	[23]
GMb	$R750/R700$	Total chlorophyll content	[27]
SR680	$R800/R680$	Total chlorophyll content	[22]
TCARI/OSAVI	$3(R700 - R670)\{-[0.2(R700 - R550)R700]/R670\}/\{[1.16(R800 - R670)]/(R800+R670 + 0.16)\}$	Total chlorophyll content	[28]
AMEC index	$(R_{swir} - R_{green})^2/R_{swir} + R_{green}$	Infestation detection	[11]
Cartenoids			
PRI	$(R531 - R570)/(R531 + R570)$	Cartenoids	[29]
Water content			
WBI	$R970/R900$	Canopy water content	[30]
Anthoyanins			
Red/green	$\sum_{i=600}^{699} R_i / \sum_{i=500}^{599} R_i$	Anthoyanins	[22]

In this study, the coupled PROSPECT[32] and SAIL[33] models, known as the PROSAIL model, were used. Three variables were tested, while the remaining parameters were fixed at typical values adapted from our earlier study.[31] Since we used a Look Up Table (LUT) approach, the first variable, leaf chlorophyll content, was fixed within a range of $25 - 105\ \mu gcm^{-2}$, the parameter N, which is the leaf structural parameter from PROSPECT, was set at 0.5–4 (unitless), and LAI, the ratio of half total leaf size projected onto the shadow was set at 5–100 (unitless). The merit function was calculated based on a least-squares function implementing CI as a weighting. CI was used for band selection, in which we attempted to select those bands sharing a higher sensitivity to the targeted parameters. For a similar reason, we used CI here again as a weighting in the inversion to focus on the more highly sensitive bands.

After obtaining the leaf chlorophyll content values, they are classified into three potential risk levels of infestation, using threshold values derived from the model training. Those highly likely to be infested have a threshold of 2, a medium likelihood assigned a value of 1, and the remainder has a value of 0, including uncertain or unknown cases.

3.2.4 *Spatial pattern feature: longitudinal profile for tree crown delineation*

Spatial information within a tree crown is useful in evaluating early leaf loss and shape change of the crown. The textural or structural anomalies in an infested crown during the early infestation stage are difficult to detect using traditional statistical textural analysis.[8] Instead, in this study, a longitudinal profile along the solar principal plane, end to end on the tree crown passing through the tree center, as developed by Zhang and Hu, was adopted.[17] A sample longitudinal profile is given in Fig. 4. In this figure, the tree crown boundary has already been segmented out, and the longitudinal profile is the line made up of digital values (e.g., reflectance) along the sun's illumination direction. To further reduce potential spectral variation and to minimize uncertainties in the GE image, the normalized relative differential values were used, wherein the differences in neighboring pixel values were divided by the average of all spectra within the crown. The

numbers of significant drops in these longitudinal profiles were used to determine the possible infestation stage.

In this study, we used the following criteria to determine the infestation stage: crossing from the center of the tree, if there was at least one major drop in values on both sides, it was considered a low infestation state, with a score of 0.5; if there were at least four drops on both sides combined, it was classified as medium infestation, with a score of 1.5; and any case with more than six such drops was considered highly infested with a score of 2.5. The necessary size and significance of these drops were determined by a threshold value derived from training.

3.2.5 *Distance from nearby known state trees from prior knowledge*

As stated in BenDor et al.[15] and Herms et al.,[16] the potential radius of infestation spread is 10–15 miles, or 16–24 km. Hence, in this study, the infestation likelihood was evaluated by considering the distance of the target tree from the known infested trees within a radius of 24 km. For calculating the distance constants, two different cases were considered in this study, street and natural park trees. Street trees were normally planted at linear and fixed intervals. Therefore, linear functions were formed and each street may be considered as a linear line, which starts and ends with known infested or uninfested trees. All trees of unknown status in between could be assigned scores of infestation likelihood. If a tree is covered by multiple functions and receives conflicting information from different linear functions, then an average is calculated

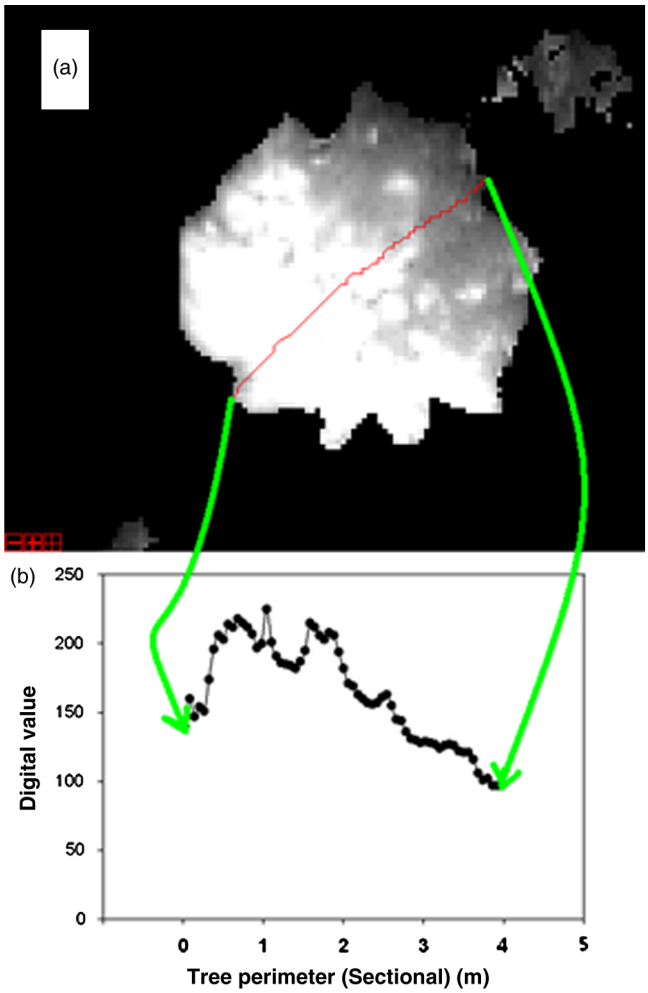

Fig. 4 A sample longitudinal profile. The tree crown is segmented out and the profile is a line made up of digital values (e.g., reflectance) along the sun's illumination direction, end to end through the tree center.

and used. For naturally occurring and park trees, a normal distribution function was formed at the location of the trees of known status, and the standard deviation was set to 20 m. This value was considered as a constant in the final evaluation function.

3.3 *Information Conversion and Final Evaluation*

3.3.1 *Training and conversion of indices and parameters*

We randomly selected half of the infested trees to obtain average values for thresholds and to generate regression functions. We used third-order polynomial functions for the VIs, and a threshold values approach for the other parameters. The procedure was repeated multiple times to ensure consistency. Each parameter was then assigned a numerical score between 0 and 5, where 0 is healthy and uninfested and 5 represents a highly infested state.

3.3.2 *Final evaluation*

After conversion, all information was assembled into the prediction function for calculating the final score, which, as above, also ranged from 0 to 5, with 0 being healthy and 5 being highly infested. The infestation score was calculated as a weighted sum of all the parameter scores related to EAB infestation. The weight for each parameter was determined based on its associated uncertainties. The final score can thus be expressed as follows:

$$\text{score} = \frac{\sum \omega_i P_i}{N \sum \omega_i} + C, \tag{2}$$

where ω_i is the weight of i'th parameter, P_i is the score of the i'th parameter, N is the dimension of the information layers, or the number of parameters used, and C represents the prior knowledge constants, which is mainly dedicated to distance information derived from prior knowledge, such as the distance from trees with known infestation states.

4 Results and Discussions

The results of this study are diverse and wide ranging, and are reported in the following sections with their associated discussions. Section 4.1 presents the prior processing studies in ground level hyperspectral signature separability assessment and contribution sensitivity analysis. Section 4.2 details the performance evaluation for the selection of the final three vegetation indices, the final vegetation indices, and their score conversion functions. Section 4.3 outlines the leaf chlorophyll content retrieval results from the PROSAIL model inversion and its score settings. Section 4.4 presents the longitudinal profile tree crown degradation study and its score settings. Section 4.5 illustrates a sample tree calculation throughout the procedural workflow. Finally, Sec. 4.6 presents the final ground validation result and the color-coded map of the estimated ash tree health.

4.1 *Prior-Processing Studies*

4.1.1 *Ground level hyperspectral signature separability assessment*

The ground level hyperspectral signature separability assessment is critical in spectral-based studies. As mentioned earlier, if no significant differences are observed between high detail (e.g., ground level) spectra of trees at different stages of infestation, then the likelihood of low detail (e.g., canopy level) spectra exhibiting sufficient differences for classification is very low. A set of sample reflectances for ash leaves from trees at different stages of infestation are shown in Fig. 5. The figure shows that there are clear differences in leaf-level hyperspectral measurements throughout the spectral range. Therefore, it may be expected that information retrieved from spectral data, particularly entire spectra, will be informative.

4.1.2 *Contribution sensitivity analysis and performance evaluation for selection of vegetation indices*

To determine bands that are potentially more successful in the VI study, the leaf-level contribution sensitivity analysis was calculated based on the PROSPECT model, which was used to evaluate the performance of various vegetation indices. The contribution analysis calculates a CI that provides quantitative measurements of the "contribution" of each observation to the final retrieval of the targeted parameter. Therefore, a higher CI represents a higher correlation. As a result, those bands with high CI should be preferred during VI band selection. We calculated a series of CI values for different leaf chlorophyll contents from the simulation studies of PROSPECT, which are shown in Fig. 6. The outcomes of these CIs show strong agreement with the findings in Sims et al.[22] and Pontius et al.[9] Figure 6 illustrates the following: (1) there are two regions, 500–600 and 700–750 nm, showing significant increase with the increase in chlorophyll content; (2) at wavelengths of less than 500 nm, the CI only changes for cases with low-chlorophyll content, typically less than 35 μgcm^{-2}; and (3) the region from 600–700 nm has a moderate effect on the CI and decreases with increasing chlorophyll content. From this, it may be expected that vegetation indices within the ranges of 500–600 and 700–800 nm will perform better than indices spanning other ranges, such as PRI, SR680 (Table 2).

4.2 *Vegetation Indices and Their Infestation Level Scores*

Besides the CI preference, the final selection of vegetation indices considered the index performance in each infestation state. We used information from trees of a known infestation state to compare the performances of the various vegetation indices, as presented in Fig. 7. In this figure, all VI values were normalized by their averages. Greater variation between different infestation stages was considered as representative of a good indicator. As such, PRI and CSc were clearly the two highest rankings, while NPQI, SR680, and TCARI/OSAVI all showed a similar level of performance. TCARI/OSAVI incorporated prefixed coefficients, which may need to be calibrated for different data sets, which is a relatively complex process. NPQI was less preferred by the CI measurements, but the bands of this index are in a different region of the spectra. The use of NPQI would increase the stability of this study, and hence it was chosen, but is given half the weight of CSc and PRI in the final function.

After examining each of the final selected vegetation indices, we found they are in good agreement with previous reports in the literature:

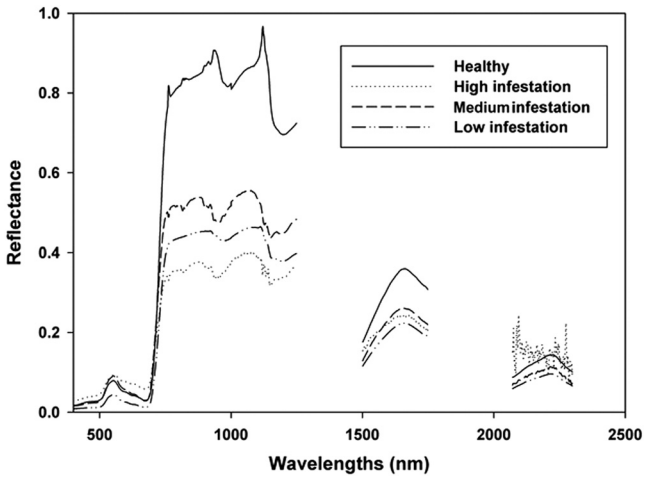

Fig. 5 Average ASD measurements of leaves (each from 19–21 samples) from ash trees that are healthy, and which have low, medium, and highly infested states. There are clear differences between the different health states throughout the entire spectral range. However, we also note that the differences in different bands are not equal and some signals may be caused by noise.

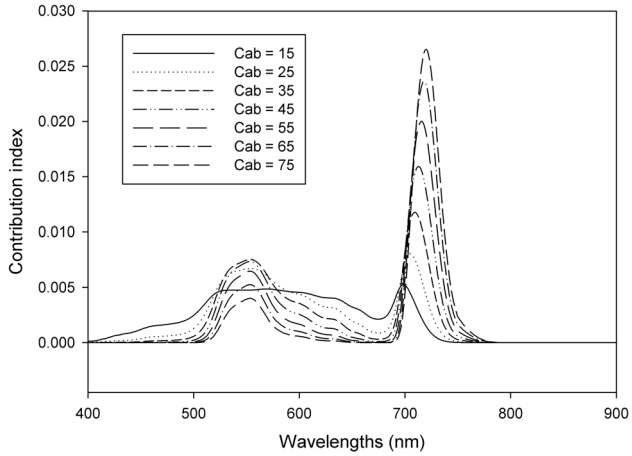

Fig. 6 Calculated contribution index based on a PROSPECT simulation at different leaf chlorophyll contents. Leaf chlorophyll content varies from 15 to 75 μgcm^{-2}. The other parameters used by PROSPECT were fixed, with at leaf structure parameter $(N[]) = 1$ to 4, equivalent water thickness $(\mathrm{Cw[cm]}) = 0.025$, and leaf dry matter per area $(\mu\mathrm{gcm}^{-2}) = 0.005$.

1) CSc, originally from Carter[30] was derived from eight different stress agents across six different plant species in laboratory studies. It was the best performing VI in our study, which agreed with the results of Pontius et al.[9]
2) PRI, which was originally derived by Gamon et al.[34] to estimate rapid changes in the relative levels of xanthophyll cycle pigments, thus serving as an estimation of photosynthetic light use efficiency,[21] which also performed well in this study, as expected.
3) The stress index NPQI,[21] which showed a positive correlation between EAB infestation and tree stress in Maire et al.,[19] but was reported as less effective in Pontius et al.[9] presented moderate performance in our analyses. However, it had relatively fewer data uncertainties than other indices, excepting CSc and PRI. We used it in the final estimation, but half weighted it compared to other indices.

All results from the vegetation indices were converted into normalized scores ranging from 0 to 5. Their functions were derived from correlations with the data of known state trees using third-order polynomial or step linear functions. The CSc, PRI, and NQPI conversion functions can be respectively given as

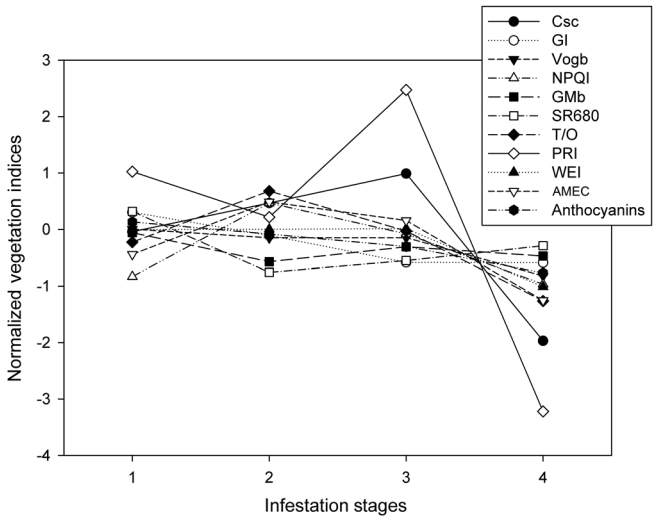

Fig. 7 Performance of selected vegetation indices. T/O indicates the TCARI/OSAVI index and AMEC is the index proposed by AMEC. For the infestation stages, 1 is healthy, 2 represents low infestation, 3 is medium infestation, and 4 represents high infestation.

$$
\begin{aligned}
P &= -0.58\mathrm{Csc}^3 + 3.503\mathrm{Csc}^2 - 5.923\mathrm{Csc} + 2.964 \\
P &= -1.835\mathrm{PRI}^3 + 12.543\mathrm{PRI}^2 - 25.590\mathrm{PRI} + 5.906 \\
P &= \begin{cases} -6.739\mathrm{NPQI} + 3.448 \ (\mathrm{NPQI} \geq 0.5) \\ -4.099\mathrm{NPQI} + 1.578 \ (\mathrm{NPQI} \leq 0.5) \end{cases}.
\end{aligned} \tag{3}
$$

Vegetation indices have provided a convenient and robust methodology for EAB detection. However, calibrations are still required for each data set. Some indices, such as NQPI, suffer from relatively low-separation performance, and thus should be utilized with reduced weighting. We also compared pixel-based spectra against the tree crown average and found the shadow effect to cause major spectral variations in near infrared (NIR) band-related indices. This also corroborates the fact that the spectral-based approach alone cannot sufficiently confirm the degree of EAB infestation.

4.3 *Retrieved Leaf Chlorophyll Content and its Infestation State Score*

The absolute leaf chlorophyll content values were obtained by PROSAIL model inversion with a CI implemented LUT approach. Hence, the hyperspectral spectra derived from the hyperspectral imagery were compared to the PROSAIL simulated spectra with CI as a weighting in the LUT inversion approach. In this study, we found that the leaf structure parameter N, ranging 1.7–2, and LAI, returning values of 3 and above, were slightly higher than reported elsewhere in the literature (e.g., Refs. 35, 36, and 37). However, they are still considered to be within a reasonable range, as most of the ash trees in our study sites were mature. We found that leaf chlorophyll content retrieved from physical models has a much higher variation than the vegetation indices. However, relative decreases in leaf chlorophyll content compared with neighboring healthy trees were found in both approaches. As a result, we did not use a regression function as in the case of the vegetation indices, and instead employed a more general classification. We used two threshold values to separate the results into three classes. The threshold values used in this study were 65 and 37 $\mu\mathrm{gcm}^{-2}$, meaning that if the leaf chlorophyll content is greater than 65 $\mu\mathrm{gcm}^{-2}$, then a score of 0 is assigned, if the value is between 37 and 64 $\mu\mathrm{gcm}^{-2}$, it receives a score of 1, and if it is less than 36 $\mu\mathrm{gcm}^{-2}$, then the score is 2.

Physical model inversion is a time-consuming process, and is prone to observation noise, thus it has not been adopted in other ash EAB infestation studies.[11] In this study, however, since it is an object-oriented approach, average tree spectra were used instead of individual pixels, significantly reducing computation time and increasing the signal-to-noise ratio. More importantly, the model inversion approach evaluates the overall spectra, taking into consideration the quality of observations, including uncertainties, and providing the absolute values of target parameters. Although the parameter we calculated in this way is leaf chlorophyll content, it still can be considered as a different approach and a fresh perspective compared to vegetation indices method, which focuses on only a few bands. We implemented CI as a weighting in the merit function to improve the retrieval accuracy. Hence, retrieving leaf chlorophyll content through physical model inversion can provide important information to be used in the final evaluation. Ideally, we seek consistent results with regard to leaf chlorophyll changes from both methods. We did confirm that the infested trees have lower leaf chlorophyll content than nearby healthy ones; however, the model inversion approach suffers from high saturation rates, which means it is fairly easy for results to converge on to the minimum or maximum values of the preset range.

4.4 *Spatial Patterns and their Infestation State Scores*

We used a normalized differential value of 0.2 as the threshold value to determine major declines in the tree crown. We found that the normalized differential longitudinal profiles show better performance with regard to resistance to uncertainties, and partially to background and shadow effects. This method simply involves counting the number of significant drops across the tree top crown profile. In Fig. 8, all trees are resampled to 30 pixels in length to illustrate them together as a normalized size. This conversion ensures low levels of infestation can be detected, while highly

infested interpretations should require further supportive evidence. We confirmed that the spatial patterning within the tree crown is a strong indicator of infestation, as previously reported.[5,13]

Figure 8 illustrates some typical examples for trees of different infestation stages. The healthy, or likely healthy, trees have a much smoother crown curve due to the healthy growing conditions, whereas the infested tree shows a clear alternating gap pattern. This pattern is due to the leaf decline within the crown, which has been reported in Pontius et al.[9] However, spatial pattern can be normally detected when the infestation is in the dieback stage. It is necessary to determine the optimal spectral band or VI from which this pattern is to be obtained, as it is also prone to variations in solar illumination conditions.

4.5 *Complete Sample Calculation Using the Operational Workflow*

Let us perform a simple but complete calculation step by step to illustrate the operational workflow introduced in this study. After segmenting each individual tree, regardless of whether this is done automatically or manually, then a spectral profile, like that shown in Fig. 9, can be obtained. These spectra can be significantly altered if unnecessary pixels are included, and care must be taken to include all pixels in the sunlit portion of the tree crown, even the dark pixels.

By calculating the selected VI from the spectrum in Fig. 9 and applying the expression functions listed in Eq. 3, we find that Csc = 0.129, giving a score of 2.25, PRI = −0.241, which gives a core of 0.48; and NPQI is −0.201, with a score of 2.398. If this spectrum is input into the PROSAIL model, with three variables and the remaining parameters set to the values mentioned in the method section, then the inversion result, giving leaf chlorophyll content, is 65. Therefore, according to our threshold values, the score from model inversion is 0.

From Fig. 10, it can be seen that there is only one major decline that crosses our threshold value. Therefore, the assigned score from the spatial pattern longitudinal profile is 0.5.

With regard to the distance from trees of known infestation state, this tree has a known medium infested tree nearby along the same street, and is spaced four trees away from it, while the nearest known healthy tree is across the street. These relationships are illustrated in Fig. 11, in which M, H, and X indicate the trees known to be medium infested, healthy, and the target tree, respectively. The target tree X is the fifth tree from the infested tree, and there is no tree between X and H; therefore, H is considered to be the sixth tree.

The final score for this individual tree is given as the solution from a linear function $y = ax + b$, with the two known points (1,2.5) and (6,0) at $x = 5$, which equates to 0.5.

Finally, the final estimation score of this individual tree, considering the score values and weightings, is given as

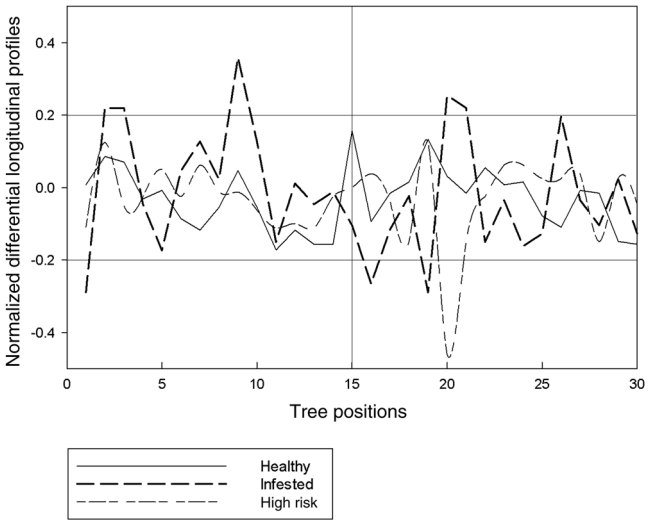

Fig. 8 Normalized differential longitudinal profiles for different representative trees. Any declines that cross the two threshold lines are considered significant. The number of significant drops is used to form the infestation estimation.

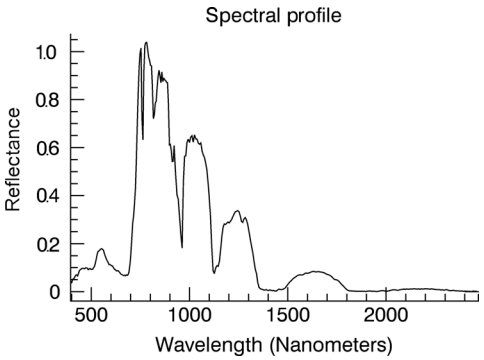

Fig. 9 A typical hyperspectral profile from the sunlit portion of the tree crown.

$$\text{score} = \frac{1 * 2.250 + 1 * 0.480 + 0.5 * 2.398 + 1 * 0.500}{3 * 3.5} + 0.500 = 0.969. \quad (4)$$

Since the score is less than 1, this tree is graded as currently healthy. Considering the typical green spectral signature attributes of a healthy living tree, the green peak, made by the green band being higher than both the blue and the red bands, and the red edge, from the NIR band being much higher than the red band, resulting in a high NDVI value, are both clearly presented. In addition, the nearby neighboring tree is currently known to be healthy. Thus, from raw observations, this tree is most likely currently healthy, which supports our calculation result. However, since there is an infested tree within the spread radius, this tree is potentially in high risk of future infestation. This is not in conflict with our score findings, since all of southern Ontario has been marked as a high risk zone. The score simply reflects the health state of the tree at the time of data acquisition.

4.6 *Map of Estimated Ash Tree Health*

Using this workflow to calculate infestation scores for all trees within our study area, the final results are collected and illustrated with ENVI (Exelis, Visual Information Solution, Boulder, Colorado, USA). The map of final results is shown in Fig. 12, where different colors indicate different estimated levels of infestation. A comparison of 40 infested trees and 40 healthy trees for assessment of estimation accuracy is shown in Table 3. Columns show the predicted result (Pd), and the rows list data from ground truthing. The overall accuracy, including healthy trees, is 70%, and 62.5% when healthy trees are excluded. The healthy tree omission error is 22.5%, and commission error is 18.5%.

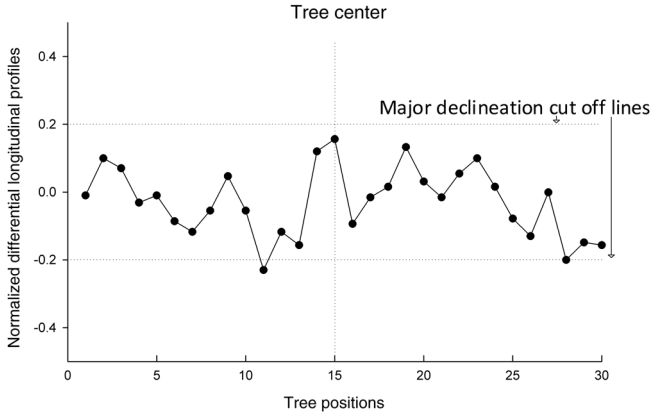

Fig. 10 Normalized differential longitudinal profile for the same tree crown as in Fig 9. The threshold value in this study is 0.2, which means any declines that have an absolute value greater than 0.2, either positive or negative is considered major and significant.

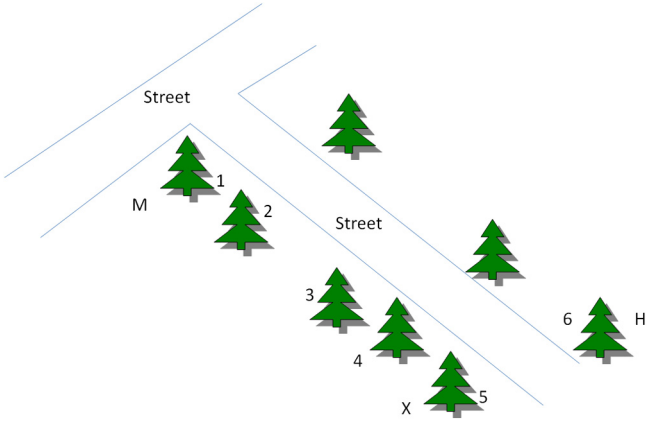

Fig. 11 A sample street scenario for distance constant calculation. M is a medium infested tree, H is a known healthy tree, and X is the targeted tree. The numbers indicate the order of trees involved in this calculation.

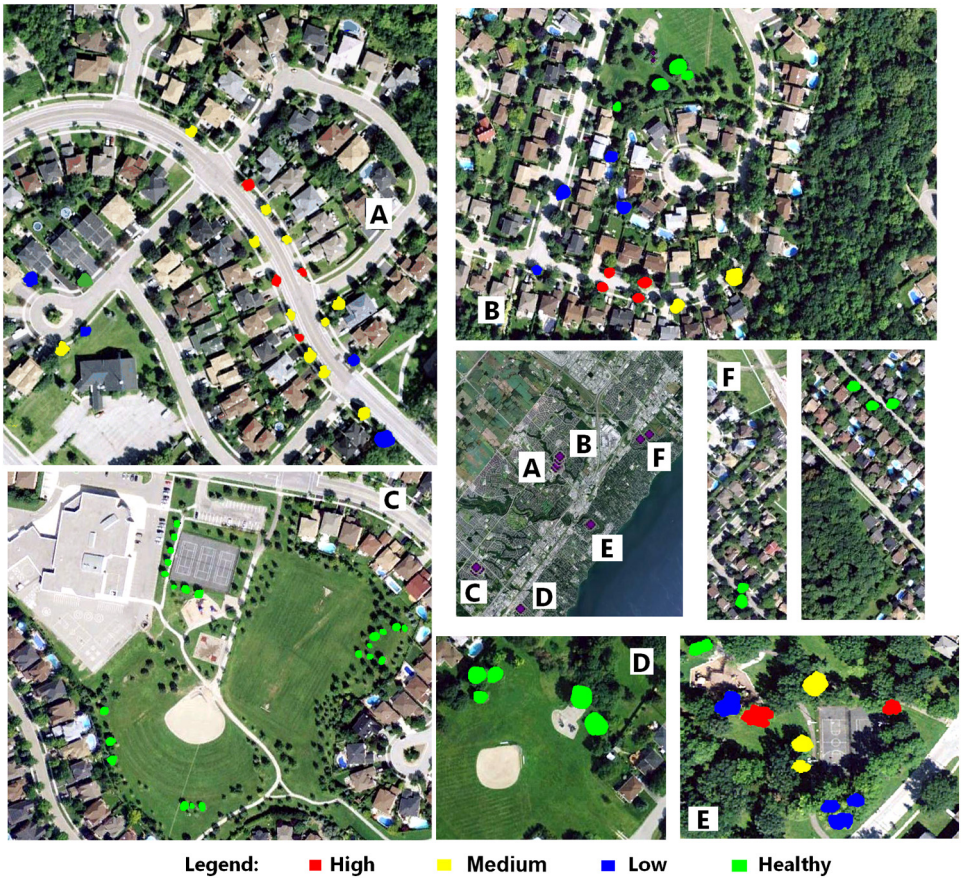

Fig. 12 Color-coded result map, where red indicates a score of 4 or higher, and highly infested trees; yellow has a score of 2.5 to 4 and represents a medium infested tree; blue, with a score of 1 to 2.5, shows low or early infested trees; and green indicates a score of less than 1 and a current healthy tree.

5 Conclusions and Future Work

This study proposed an object-oriented approach to detecting early infestation of EAB from multisourced data, such as hyperspectral and high spatial resolution imagery. We derived different types of information from these data, which can be used together to detect EAB infestation. All information derived in this way is utilized in a weighted prediction function, in which all data are normalized into scores of 0–5.

Table 3 Accuracy assessment of for estimated and ground observed tree health states for 80 sample trees.

Predicted/ground	High	Medium	Low	Healthy
High	10	0	0	0
Medium	2	8	1	5
Low	3	2	7	4
Healthy	1	1	5	31

The hypothesis of correlation between general stress and EAB infestation is confirmed. EAB infested trees show variations in leaf chlorophyll content, predominantly a reduction to, for example, less than 37 μgcm^{-2}, as a symptom of stress. From an observational point of view, stressed ash trees are more likely to be infested by EAB, and therefore stress-associated variations in leaf chlorophyll content, normally to values lower than that found in nearby trees of the same species, are considered as the primary indicator in detecting early infestations in ash trees. However, despite drops in leaf chlorophyll content having over 85% correlation with EAB infestation, the technique suffers from high omission errors. Thus, while the infested trees most likely have reduced leaf chlorophyll content, low-leaf chlorophyll content does not always confirm an infestation.

We found that vegetation indices, leaf chlorophyll content, longitudinal profiles, and paper documentation all have the potential to detect EAB infestation. While vegetation indices are the most convenient and robust methods, leaf chlorophyll content and longitudinal profiles may provide more precise results, but at the cost of computation time and the risk of considerable variation. Paper documents are considered as prior knowledge in this study, and represent the most reliable information. The synergy of web GIS and remote sensing techniques would be an interesting direction for future investigation.

Despite the weaknesses and limitations of GE's imagery, it is easy to access, free of charge, offers multiple years of availability, and in some cases provides better spatial resolution. These benefits make it a viable option for resource management and planning by governments and other agencies. Most importantly, GE imagery and the relative profile approach can be easily adopted and effectively employed by nonspecialists.

In this study, all segmentations underwent manual corrections, which ensured that the initial tree boundary and location information were accurate. However, manual segmentation and validation are very labor intensive and time consuming, and for this scheme to be applied over a larger area, automatic segmentation would be essential, which implicitly suggests that errors from segmentation could propagate and directly affect the accuracy of the final evaluation.

The PROSAIL model was chosen due to its simplicity, reliability, robustness, and effectiveness. In this study, PROSAIL served as a demonstration, which had been validated and reported with successful collaboration with CI-based optimization in Zhang et al.[31] This does not exclude other models from being adopted in this approach, and a comparison of the effectiveness and accuracy of different models would be an interesting future investigation.

The final accuracy validation is also limited by the data sources available, particularly in the number of confirmed infested trees. This and future studies could therefore be improved with further investigations and explorations.

This study presents a demonstration of a proof of concept and a prototype workflow. The vegetation indices chosen and the models used in parameter retrieval processing are not fixed, and may require future investigation for possible improvements. Other spatial information, such as texture, may also be considered for potential integration and expansion of the calculation.

Acknowledgments

The authors would like to thank Ziya He and Meaghan Eastwood of Toronto and Region Conservation for their contributions in the field and in consulting this research, and the town of Oakville, AMEC Inc., and the York University library for kindly supplying the data

and technical support. The authors are grateful for financial support provided by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

1. K. Maloney, J. Boughton, and N. Schneeberger, *Emerald Ash Borer 2006 Brief*, USDA Forest Service, Northeastern Area, State and Private Forestry, Newtown Square, Pennsylvania (2006).
2. T. D. Sydnor, M. Bumgardner, and A. Todd, "The potential economic impacts of emerald ash borer on Ohio US communities," *Int. Soc. Arboriculture* **1**(33), 45–54 (2007).
3. J. S. Gould et al., *Emerald Ash Borer, Agrilus Planipennis (Fairmarire), Biological Control Release and Recovery Guideline*, United States Department of Agriculture (2013).
4. K. F. Kovacs, "Cost of potential emerald ash borer damage in U.S. communities," *Ecol. Economic.* **69**(3), 569–578 (2010).
5. J. R. Rose, *Landowners Guide for Woodlot Threaded by Emerald Ash Borer*, Ontario Ministry of Natural Resource (2010).
6. A. C. Anulewicz, D. G. McCullough, and D. L. Cappaert, "Emerald ash borer (Agrilus planipennis) density and canopy dieback in three North American ash species," *AUF* **5**(33), 338–349 (2007).
7. K. L. Ryall, J. G. Fidgen, and J. J. Turgeon, "Detectability of the emerald urban tree by using branch samples," *Environ. Entomol.* **40**(3), 679–688 (2011).
8. D. G. McCullough and A. S. Katovich, *Emerald Ash Borer. Pest Alert*, Newtown Square, PA:USDA Forest Service NA-PR-02-04 (2004).
9. J. Pontius et al., "Ash decline assessment in emerald ash borer-infested regions: a test of tree-level, hyperspectral technologies," *Remote Sens. Environ.* **112**, 2665–2676 (2008).
10. J. S. Souci, I. Hanou, and D. Puchalski, "High resolution remote sensing analysis for early detection and response planning for emerald ash borer," *Photogramm. Eng. Remote Sens.* **75**(8), 905–911 (2009).
11. I. Hanou, "Town of Oakville hyperspectral EAB analysis," Town of Oakville government document, http://www.oakville.ca/assets/general%20-%20residents/eab-hyperspectral.pdf (2010).
12. D. Smitley, T. Davis, and E. Rebek, "Progression of ash canopy thinning and dieback outward from the initial infestation of emerald ash borer (Coleoptera: Buprestidae) in southeastern Michigan," *J. Economic. Entormol.* **101**, 1643–1650 (2008).
13. P. Cosmopoulos and D. J. King, "Temporal analysis of forest structural condition at an acid mine site using multispectral digital camera imagery," *Int. J. Remote Sens.* **12**(25), 2259–2275 (2008).
14. K. Omari et al., "Retrieval of forest canopy parameters by inversion of the proflair leaf canopy reflectance model using the LUT approach," *IEEE J. Sel. Topics Appl. Earth Observ.* **6**(2), 715–723 (2013).
15. T. K. BenDor et al., "Modeling the spread of the emerald ash borer," *Ecol. Model.* **197**(1), 222–236 (2006).
16. D. A. Herms et al., *Insecticide Options For Protecting Ash Trees From Emerald Ash Borer*, North Central IPM Center Bulletin, Canada (2009).
17. K. Zhang and B. Hu, "Individual urban tree species classification using very high spatial resolution airborne multi-spectral imagery using longitudinal profiles," *Remote Sens.* **4**, 1741–1757 (2012).
18. L. Jing et al., "Automated tree crown delineation from imagery based on morphological techniques," in *Proc. 35th Int. Symposium on Remote Sensing of Environment*, pp. 22–26, Beijing (2013).
19. G. le Maire, C. Francois, and E. Dufrene, "Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements," *Remote Sens. Environ.* **89**, 1–28 (2004).
20. S. Thenkabail, J. G. Lyon, and A. Huete, *Hyperspectral Remote Sensing of Vegetation*, CRC Press, Taylor and Francis Group (2012).

21. Y. Zhang et al., "Leaf chlorophyll content retrieval from airborne hyperspectral remote sensing imagery," *Remote Sens. Environ.* **112**, 3234–3247 (2008).
22. D. A. Sims and J. A. Gamon, "Relationships between leaf pigment content and spectral reflectance across a wide range of species leaf structures and development stages," *Remote Sens. Environ.* **81**, 337–354 (2002).
23. J. D. Barnes et al., "A reappraisal of the use of DMSO for extraction and determination of chlorophylls a and b in lichens and higher plants," *Environ. Exp. Bot.* **32**(2), 85–100 (1992).
24. G. A. Carter and R. L. Miller, "Early detection of plant stress by digital imaging within narrow stress-sensitive wavebands," *Remote Sens. Environ.* **50**, 295–302 (1994).
25. A. M. Smith et al., "Leaf area index from CHRIS satellite and application in plant yield estimation," in *Proc. 26th Canadian Symposium on Remote Sensing*, Wolfville, Nova Scotia (2005).
26. J. E. Vogelmann, B. N. Rock, and D. M. Moss, "Red edge spectral measurements from sugar maple leaves," *Int. J. Remote Sens.* **14**, 1563–1575 (1993).
27. A. A. Gitelson and M. N. Merzlyak, "Quantitative estimation of chlorophyll-a using reflectance spectra experiments with autumn chestnut and maple leaves," *J. Photochem. Phytobiol.* **22**, 247–252 (1994).
28. D. Haboudane et al., "Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture," *Remote Sens. Environ.* **81**, 416–426 (2002).
29. M. N. Merzlyak et al., "Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening," *Physiol. Plantarum* **106**, 135–141 (1999).
30. G. A. Carter, "Ratio of leaf reflectance in narrow wavebands as indicators of plants stress," *J. Int. Remote Sens.* **15**, 697–703 (1994).
31. K. Zhang et al., "Improving the retrieval of the biophysical parameters of vegetation canopies by using the contribution index," *Canadian J. Remote Sens.* **37**(6), 643–652 (2012).
32. S. Jacquemoud and F. Baret, "Prospect: a model of leaf optical properties," *Remote Sens. Environ.* **34**(2), 75–91 (1990).
33. W. Verhoef, "Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model," *Remote Sens. Environ.* **16**, 125–141 (1984).
34. J. A. Gamon, J. Penuelas, and C. B. Field, "A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency," *Remote Sens. Environ.* **41**, 35–44 (1992).
35. P. Zarco-Tejada et al., "Scaling-up and model inversion methods with narrow-band optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data," *IEEE Trans. Geosci. Remote Sens.* **39**, 1491–1507 (2001).
36. D. Haboudane et al., "Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture," *Remote Sens. Environ.* **90**, 337–352 (2004).
37. S. Jacquemoud et al., "PROSPECT + SAIL models: a review of use for vegetation characterization," *Remote Sens. Environ.* **113**, S55–SS66 (2009).

Kongwen Zhang is a PhD student at York University. He received his BASc and MSc degrees in geomatics engineering and Earth science from York University in 2005 and 2007, respectively.

Baoxin Hu is an associate professor at York University. She received her BSc degree from Tianjin University, China, an MSc degree from Chinese Academy of Sciences, and a PhD degree from Boston University, USA, in 1992, 1994, and 1998, respectively.

Justin Robinson is a research assistant at Selkirk Geospatial Research Centre. He received his environmental technology dip. from Camosun College in 2010 and an adv. dip. in GIS from Selkirk College in 2011.