Sentinel-2 accurately maps green-attack stage of European spruce bark beetle (*Ips typographus*, L.) compared with Landsat-8

Haidi Abdullah¹, Andrew K. Skidmore¹,², Roshanak Darvishzadeh¹ & Marco Heurich³,⁴

¹Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands
²Department of Environmental Science, Macquarie University, NSW, 2106, Australia
³Department of Conservation and Research, Bavarian Forest National Park, 94481, Grafenau, Germany
⁴Chair of Wildlife Ecology and Wildlife Management, University of Freiburg, Tennenbacher Straße 4, Freiburg, Germany

Keywords
bark beetle (*Ips typographus* L.), green attack, Landsat-8, Norway spruce, Sentinel-2, spectral vegetation indices

Abstract
Natural disturbances induced by insect outbreaks have increased in forest ecosystems over the past decades. To minimize economic loss and prevent a mass outbreak, early detection of bark beetle green attack – a period when trees have yet to show visual signs of infestation stress – is therefore crucial to effective and timely forest management. In this study, we evaluated the ability of spectral vegetation indices extracted from Landsat-8 and Sentinel-2 imagery to map bark beetle green attack using principal component analysis (PCA) and partial least square discriminate analysis (PLS-DA). A recent infestation map produced through visual interpretation of high-resolution aerial photographs validated the final infestation output maps. Leaf spectral measurements alongside total chlorophyll and nitrogen concentration, leaf water content and leaf dry matter content were measured to assess the impact of bark beetle green attack on foliar properties. We observed that the majority of spectral vegetation indices (SVIs) calculated from Sentinel-2, particularly red-edge dependent indices (NDRE 2 and 3) and water-related indices (SR-SWIR, NDWI, DSWI and LWCI), were able to discriminate healthy from infested plots. In contrast, only the water-related indices (NDWI, DSWI and RDI) from Landsat-8 were able to discriminate between healthy and infested plots efficiently. The total number of pixels identified as harboring a green attack that matched with ground truth data (aerial photography) was higher for Sentinel-2 (67%) than for Landsat-8 (36%) SVIs, indicating the elevated sensitivity of Sentinel-2 imagery to changes induced by bark beetle green attack. We also determined that foliar chlorophyll and leaf water content were significantly higher \((P < 0.05)\) in healthy trees than in green-attacked trees. Our study highlights the potential of Sentinel-2 data for the early detection of bark beetle infestations and the production of reliable infestation maps at the green-attack stage.
Introduction

Insect infestations, as by the European spruce bark beetle (*Ips typographus* L.), form the main disturbance events in European forests as they destroy more forested areas than all other natural disturbances together. Historically, storms and snow breakage events in Europe and North America have provided a large surplus of suitable breeding material for bark beetles, leading to outbreaks (Giunta et al. 2016; Seidl et al. 2017). Recent outbreaks are different, however, with climate change (more drought and severe storm events) appearing to be the initiating factor (Netherer et al. 2015; Thom and Seidl 2016; Marini et al. 2017). Increasing temperatures may result in an increased incidence of drought, possibly affecting tree health. Drought facilitates beetle outbreaks by stressing trees and increasing the frequency and severity of bark beetle outbreaks (Bentz et al. 2010; Filchev 2012; Faccioli and Bernardinelli 2014; Thom and Seidl 2016). A warmer climate may increase storm frequencies, and severity of wind throws, thus providing more breeding material for bark beetles (Marini et al. 2017). Furthermore, as the development of the bark beetle is dependent on temperature, an increase in temperature may thus lead to an increase in the beetle population size, due to their eggs hatching and developing into adults sooner (Wermelinger 2004; Bentz et al. 2010). In recent decades, therefore, increasing disturbances due to insect outbreaks have been widely documented across different parts of the world (Wulder et al. 2006; Lausch et al. 2013a; Seidl et al. 2014).

Timely, accurate and cost-effective information is needed to mitigate and control bark beetle outbreaks, to guide forest managers in identifying areas infested by beetles, as well as to define the timing of bark beetle control activities (Wulder et al. 2009). To obtain this information, the operation survey needs to take place at the time when the infestation at its early stage (the green attack). The green-attack stage is the first interaction between the beetle and the host tree, and occurs when the host is being colonized by the bark beetle. During this stage, the infested tree is still physiologically green and very much alive, although exhibiting stress in the near infrared (invisible to the human eye) (Niemann and Visintini 2005). Furthermore, during the green infestation, the newly hatched generation of beetles is developing within the inner bark of the infested trees. Therefore, management intervention to prevent further outbreaks may involve the removal of infested trees before the new brood emerges and migrates (Wermelinger 2004; Wulder et al. 2009).

Traditionally, foresters perform field surveys to identify infested trees during the early green-attack stage using conventional survey methods (looking for sawdust). Such surveys are very laborious and therefore make screening large areas for green attack difficult. Nowadays, remote sensing provides new opportunities to detect and map bark beetle infestation. Remotely sensed data rely on spectral signatures from different regions in the electromagnetic spectrum. Unique spectral signatures have been linked to different functional and structural plant traits, such as pigments at 400–700 nm, leaf structure at 700–1100 nm and plant water content, nitrogen concentration, Leaf Area Index and Specific Leaf Area at 1100–2400 nm (Gitelson et al. 2002; Viña et al. 2011; Mahlein et al. 2013; Wang et al. 2015; Ali et al. 2016). Stress affects plant biophysical and biochemical properties and therefore affects spectral signature. For example, chlorophyll degradation and nitrogen deficiency lead to an increase in reflectance spectra in the visible region (the red and green bands in particular). As a result, this wavelength region has been widely used as a stress indicator when utilizing remote sensing data (Hendry et al. 1987). Significantly, the reflectance at the red portion of the visible region has been shown to be less sensitive to initial loss of chlorophyll content (Carter 1993). This is due to the high spectral absorption by chlorophyll in this spectral region, which saturates the red reflectance at low chlorophyll content (Jacquemoud and Baret 1990). In contrast, the reflectance at the green and red-edge wavelengths (centered in 550 nm and 700 nm, respectively) of the VIS region are more sensitive to changes in plant chlorophyll content (Eitel et al. 2011). Furthermore, the red-edge region has superior sensitivity when detecting changes that are induced in plants by stressors such as dehydration, disease and insect attack, and can, therefore, improve the early detection of plant stress (Ahern 1988; Carter and Miller 1994; Carter and Knapp 2001; Eitel et al. 2011). Moreover, the NIR and SWIR regions have been used widely to assess water content and nitrogen concentration in plants (Carol et al. 2004; Jackson 2004; Ayala-Silva and Beyl 2005; Munoz-Huerta et al. 2013).

Several studies have utilized spectral vegetation indices (SVIs) from low-to-medium resolution satellite data to study bark beetle infestation. These studies have mainly focused on the last two infestation stages (red and grey) of attacks (Franklin et al. 2003; Hais and Kuci et al. 2008; Havašová et al. 2015; Meddens et al. 2013; Wulder et al. 2006). However, for effective and proper forest management, the detection of infestation should be early enough to allow for timely intervention to minimize the outbreak. Several studies have explored the use of commercial remote sensing data such as Worldview-2 (Filchev 2012; Immnitzer and Atzberger 2014), RapidEye (Marx and der Havel 2010; Ortiz et al. 2013) and HyMAP airborne hyperspectral data (Lausch et al. 2013b) for the early detection of bark beetle attack, but with very limited success.
The SVIs were introduced to improve interpretation of vegetation signals when using remote-sensing data and can be used to measure vegetation status while minimizing solar irradiance and soil background effects (Jackson and Huete 1991; Moulin 1999; Darvishzadeh et al. 2009). In addition, the combination of different spectral bands is closely related to biophysical and biochemical properties of foliage vigor associated with plant health and can be used to detect morphological and physiological changes caused by insect outbreaks in the forest canopy (Rullan-Silva et al. 2013). Therefore, SVIs may be expected to perform better than individual spectral bands when it comes to detecting stress induced by the bark beetle. This is especially valid when using low-to-medium resolution data that include the red-edge spectral domain (Jackson and Huete 1991; Zhang et al. 2012). This study evaluates the ability of different SVIs from Sentinel-2 and Landsat-8 imagery to detect and help map bark beetle infestation at the green-attack stage. Furthermore, we also study the impact of bark beetle green attack on foliar biochemical and biophysical properties and their spectral reflectance using foliar spectral data collected from ASD FieldSpec3.

Materials and Methods

Study area and field data collection

The study area is the Bavarian Forest National Park, which is a 24,369 ha forest located in south-eastern Germany along the broader with the Czech Republic, between 13°12'9"E (longitude) and 49°3'19"N (latitude). This region is characterized as having a temperate climate with an annual precipitation of between 900 and 1800 mm, and a mean annual temperature of between 13°C and 9°C (Bässler et al. 2008). The forest is dominated by Norway spruce (Picea abies) (67%) and European beech (Fagus sylvatica) (24.5%) (Cailleret et al. 2014). Outbreaks of the bark beetle (Ips typographus L.) began in 1984 and have caused extensive disturbance to this forest. As such, the area is a suitable study site for research on bark beetle infestations and outbreaks (Heurich et al. 2010).

During June and the beginning of July 2016, an extensive campaign was conducted to collect field measurements. The study area was divided into two different strata based on their tree condition: stands with healthy trees and stands with trees freshly infested by the bark beetle. To select the healthy stands, a stratified random sampling strategy was adopted. While, for the recently infested stands, the presence of bark beetle green attack had to be confirmed in an intensive field survey by searching for piles of dry, boring dust pushed out onto the bark surface of the tree when the beetle tunnels under the bark. According to the Bavarian Forest National Park authorities and our field survey, the infestation of bark beetles in 2016 clearly indicated outbreak conditions.

To avoid mixed reflectance from healthy and green-attacked trees, only plots fully under bark beetle green attack that covered an area of 30 m × 30 m were selected. In total, 40 and 21 plots were selected in healthy and infested stands, respectively (Fig. 1). The centre of each plot was measured using Differential Global Positioning System (DGPS) Leica GPS 1200 (Leica Geosystems AG, Heerbrugg, Switzerland) with an accuracy of better than 1 m after post-processing. Each plot was designed as a 30 × 30 m square and within the plot stand characteristics including DBH, canopy cover, height and tree density were measured. At each plot, three to five trees were selected as representative and needle samples from each of these trees were collected separately. All the samples were taken from the trees’ top layer exposed to the sunlight. A crossbow was used to shoot an arrow attached to a fishing line at a branch with sunlit leaves (Ali et al. 2016).

Leaf spectral reflectance and traits including total chlorophyll, leaf water content (Cw), leaf dry matter content (LDMC), and foliar nitrogen concentration were measured for healthy and infested samples (Table 1). In the field, the concentration of chlorophyll was determined using a CCM chlorophyll meter. On average ten readings were immediately taken from each fallen branch with the CCM. The needles were then removed from the fallen branches, covered with wet pulp paper and placed in a labeled plastic zip-locked bag. They were transported to the laboratory using a portable cooling box; this was done to retard possible changes in the needles’ reflectance spectra and biochemical characteristics. The leaf directional hemispherical reflectance from 350 to 2500 nm was measured for the collected samples, using an ASD FieldSpec-3 Pro FR spectrometer equipped with an ASD RT3-32C integrating sphere (Analytical Spectral Devices, Inc. Boulder, CO, USA). Further details concerning the measurements of hemispherical needle reflectance can be found in Abdullah et al. (2018), Ali et al. (2016), Daughtry et al. (1989), and Malenovský et al. (2006). The needle samples were then dried for 72 h using an oven at 60°C, until they reached constant weight, to calculate leaf dry matter and water content (see Table 1). To determine foliar nitrogen, the dried needles were ground using mortar and pestle until they became a soft powder, after which they were passed through a 0.25 mm mesh screen. Subsequently, 2 mg of powdered leaves were transferred to a small aluminum capsule to measure the nitrogen content using an organic elemental analyser (FLASH 2000). The studied leaf traits (chlorophyll, water content, dry matter content and nitrogen concentration) from the
measured representative trees in each plot were then averaged to obtain the leaf traits at the plot level, hereafter referred to as plot level parameters.

**Satellite imagery**

Open access multispectral satellite imagery from Sentinel-2 and Landsat-8 was selected for this study. To fulfill the aims of this research, cloud-free satellite data were acquired on 8 and 10 July 2016 from Sentinel-2 and Landsat-8, respectively. Landsat-8 data were obtained from USGS Global Visualization Viewer (http://glovis.usgs.gov/), and the Sentinel-2 data were obtained from the ESA Scientific Hub (https://scihub.copernicus.eu). Both images were geometrically corrected in the Universal Transform Mercator (UTM) coordinate system and matched each other with sub-pixel accuracy. Sentinel-2 delivers high spatial (10 to 20 m) and spectral (13 bands) image data. Also, it is the first freely available optical satellite providing three spectral bands in the red-edge spectral region.

To prepare the satellite images for further analysis, a series of pre-processing was applied. First, for Landsat-8, image radiometric calibration was applied to convert the pixel values to Top-of-Atmosphere (TOA) reflectance. A MODTRAN4-based atmospheric correction software package (FLAASH) developed by the Air Force Phillips Laboratory, Hanscom AFB, and Spectral Sciences, Inc. was used to convert the TOA Reflectance to surface reflectance (Adler-Golden et al. 1999) using ENVI software. For Sentinel-2, the TOA reflectance data were corrected to surface reflectance using the SEN2COR atmospheric correction software developed by ESA (http://step.esa.int/main/third-party-plugins-2/sen2cor/).

Secondly, the bands with a similar spatial resolution were stacked. In general, six bands from Landsat-8 were used
and stacked together, namely the bands 2, 3, 4, 5, 6 and 7, which have a 30 m spatial resolution. From the Sentinel-2, the bands 2, 3, 4, and 8, with a 10 m spectral resolution, were first resampled to 20 m and then stacked with bands 5, 6, 7, 9, 12, and 13. The three bands with a 60 m spatial resolution (1, 10, 11) are mainly relevant for atmospheric corrections and were not used in this study. Finally, the spectral reflectance values of the sample plots were extracted from the Landsat-8 and Sentinel-2 scenes and used for further analysis. To synthesize and compare leaf reflectance data with canopy reflectance data (from Sentinel-2 and Landsat-8 data), the reflectance spectra collected from ASD FieldSpec-3 Pro FR spectrometer equipped with an ASD RT3-3ZC integrating sphere were simulated by convolving to the spectral resolution of Sentinel-2 and Landsat-8 using the linear interpolation method.

### Spectral vegetation indices calculation

Spectral reflectance data from Sentinel-2 and Landsat-8 images were used to calculate the most widely used vegetation indices to detect changes in plant photosynthetic activity and biochemical stress, among other forms of vegetation stress (Collins and Woodcock 1996; Eitel et al. 2006). Spectral vegetation indices include optical vegetation canopy (greenness), which is a combined property of foliar biochemical (chlorophyll, nitrogen and leaf water content) and other canopy properties (Jiang et al. 2008). To prevent using multiple copies of similar band combinations such as (red and NIR) in this study, we attempted to use statistically independent spectral indices. In addition, we employed the raw spectral bands as independent indices and only those bands that were significantly different ($P \leq 0.05$) between healthy and infested sample plots (Table 2). The selected SVIs are sensitive to stress-induced variations in chlorophyll content (VIS), biomass (NIR), and water content (SWIR). Furthermore, depending on their use, we categorized the vegetation indices into three groups: (a) chlorophyll and other pigments, (b) indices used for detection of water stress, and (c) the raw spectral bands from Sentinel-2 and Landsat-8.

### Statistical analysis

Three statistical analyses were employed in this study: one-way ANOVA, principal component analysis (PCA), and partial least square discriminant analysis (PLS-DA). ANOVA tests were performed to: (a) ascertain the effect

---

**Table 1.** The leaf traits measured for healthy and infested samples in this study. Fw, Dw and A represent fresh leaf weight (g), dry leaf weight (g), leaf dry mass per unit area (Cm) and leaf area (cm²), respectively.

<table>
<thead>
<tr>
<th>Leaf traits</th>
<th>Equation</th>
<th>Unit</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaf water mass per area (Cw)</td>
<td>$\frac{Fw - Dw}{A}$</td>
<td>mg/cm²</td>
<td>Danson et al. (1992) and Ceccato et al. (2001)</td>
</tr>
<tr>
<td>Leaf dry matter content (LDMC)</td>
<td>$\frac{Cm}{A}$</td>
<td>mg/g</td>
<td>Vile et al. (2005)</td>
</tr>
<tr>
<td>Nitrogen (%)</td>
<td>$\frac{Cm - Cw}{A}$</td>
<td>%</td>
<td>PerkinElmer 2400 CHNVO</td>
</tr>
<tr>
<td>Chlorophyll</td>
<td>$\frac{Cm - Cw}{A}$</td>
<td>mg/m²</td>
<td>CCM-300 chlorophyll content meter</td>
</tr>
</tbody>
</table>

**Table 2.** One-way ANOVA test between healthy and infested reflectance data at both leaf (simulated) and canopy level.

<table>
<thead>
<tr>
<th>Spectral bands</th>
<th>Landsat-8</th>
<th></th>
<th>Sentinel-2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Leaf level</td>
<td>Canopy level</td>
<td>Leaf level</td>
<td>Canopy level</td>
</tr>
<tr>
<td>Blue</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>Green</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td>***</td>
</tr>
<tr>
<td>Red</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td>***</td>
</tr>
<tr>
<td>Red-edge1</td>
<td>Not available</td>
<td>Not available</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Red-edge2</td>
<td>Not available</td>
<td>Not available</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Red-edge3</td>
<td>Not available</td>
<td>Not available</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>NIR</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td>***</td>
</tr>
<tr>
<td>NIR(a)</td>
<td>Not available</td>
<td>Not available</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>SWIR-1</td>
<td>*</td>
<td>*</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>SWIR-2</td>
<td>*</td>
<td>*</td>
<td>***</td>
<td>***</td>
</tr>
</tbody>
</table>

(*) Not significant, (*) Hardly significant, (**) significant, (*** Strongly significant.
of bark beetle green attack on measured leaf traits, including leaf and canopy reflectance data and (b) to evaluate whether SVIs’ between the two sample groups (healthy and infested) were significantly different. Moreover, to visually examine how well the two sample plots (healthy and infested) are separated in 2-D space based on principal component analysis, the clustering method using principal component analysis is employed in this study. As such, PCA was used to evaluate the potential for different spectral vegetation indices to differentiate between healthy and infested sample plots. PCA is an unsupervised technique and remains a popular method used to reduce the dimension of multivariate datasets and to extract features (Pearson 1901; Hotelling 1933). The unit variance and mean centering producer were used to pre-process the data. To build the PCA model, spectral vegetation indices were treated as independent variables, and the SVIs data of all 60 sample plots (21 infested and 40 healthy plots) were analyzed. PCA models were built independently for each SVIs group as follows: (a) including all SVIs (pigment, water and raw bands); (b) including only pigment-related indices; and (c) including only water-related indices. This step is important because it allows us to understand and identify the most effective spectral indices group for separating the healthy from the infested sample groups (Fig. 2).

In order to identify the key SVIs influencing the spectral separability between the healthy and infested plots, and to map bark beetle green attack, a novel method known as a Random Frog (RF) was used. The RF is a state-of-the-art variable selection algorithm, and it is a computationally efficient method using the context of the reversible jump Markov chain Monte Carlo (MCMC) technique (Green 1995). It performs a search in the model space via both fixed-dimensional and trans-dimensional moves between different models. After a pseudo-MCMC chain is calculated, this can be used to calculate a selection probability (SP) value for each variable included in the model. The key variables can be identified regarding the ranking of all variables based on the SP value. A detailed description of the calculation of RF can be found in Yun et al. (2013). To achieve this, partial least squares-discriminant analysis (PLS-DA) was employed as a modeling method in RF (Li et al. 2012). PLS-DA is a classification technique allowing for the identification of variables that improve the separation or classification between different groups (Wold et al. 2001).

In our study, SVIs with an SP value >0.50 were selected as an important variable to map bark beetle green attack. Following this, the selected SVIs from both satellites (Sentinel-2 and Landsat-8) were used to map bark beetle green attack. The box plot technique was used to display the distribution of SVI values for healthy and infested plots. The threshold values for green-attacked pixels were identified for each selected SVI. The criteria for selecting the threshold values were based on the area (value) of each SVI, which essentially characterized the difference (no overlap) between healthy areas and those infested by bark beetle green attack (Fig. 9). Consequently, a conditional decision was made using the identified threshold values from each SVI to identify those pixels in the images falling within the identified threshold range (assigned as 1) and those pixels falling outside of the threshold range (assigned as 0). For example, the following conditional decision was utilized to extract the threshold value from NDRE-2:

\[
\text{Green attacked pixels} = \begin{cases} 
1, & \text{if } 0.45 \leq \text{NDRE2} \leq 0.80 \\
0, & \text{otherwise}
\end{cases}
\]

(1)

From the conditional statement the following threshold values are extracted: if the value of NDRE-2 is higher than 0.45 and <0.80, a 1 (green attack) will be assigned to that cell location on the output raster; otherwise, a 0 (false) will be assigned on the output raster. Similar conditions were applied to the other selected SVIs (Fig. 8). Finally, to generate the final infestation map, the following equation was applied to sum the infestation maps generated from each SVIs:

\[
\text{Final infestation map} = \begin{cases} 
1, & \text{if indicesA and B and C...} = 1 \\
0, & \text{otherwise}
\end{cases}
\]

(2)

All of the above statistical analyses were carried out using MATLAB 2016b (MathWorks Inc, Natick, USA) and ArcMap (v.10.3).

**Ancillary data and accuracy assessment**

To assess how successfully the infestation maps produced in this study matched with the existing infestation areas, reference data corresponding to the locations of infested trees obtained from the Bavarian Forest National Park administration were used for validation of the output infestation maps. The flight on 11 Jun 2017 documented new dead wood (grey-attack stage) from the previous year 2016, with reference data collected from airborne color-infrared (VIS and NIR) aerial images (CIR) with 0.1 m spatial resolution. Full details about the processing and interpretation of the aerial photography in the Bavarian Forest National Park can be found (Heurich et al. 2009; Lausch et al. 2013a).

The infestation data were in the form of polygons and were rasterized twice; firstly into 20 m × 20 m grid cells to match the generated map of infestation from Sentinel-2.
SVIs, and secondly into 30 m × 30 m grid cells to match the infestation map generated from Landsat-8 SVIs (Fig. 2). As bark beetles only infest old and mature Norway spruce (Picea abies (L.) Karst), the non-spruce stands and young stands were masked using land cover data obtained from the national park administration (Dupke et al. 2017). The masked land cover was overlaid on the infestation maps generated in this study and compared with the rasterized reference data. Finally, the total number of pixels that correctly matched with the reference pixels (ground truth) were extracted and calculated (Fig. 2).

Results
Impact of bark beetle green attack on measured leaf traits
The results of the ANOVA test showed significant differences (P < 0.05) between healthy and infested samples for all measured leaf traits in this study except for nitrogen concentration at plot level (Fig. 3). At the leaf level, chlorophyll and leaf water content were significantly higher for healthy than for infested foliar. In contrast, smaller differences were observed in foliar nitrogen concentration (P < 0.04) between these two groups. Leaf dry matter content (LDMC) was significantly higher for the infested leaves. Similarly, at plot level, higher chlorophyll and water content was observed for the healthy plots (Fig. 3).

Leaf and canopy spectral variations
Figure 4 shows the difference between the mean reflectance of healthy and infested foliage. The difference is largest in the VIS region between 520 and 685 nm, the NIR (740–1130 nm) and the shortwave infrared region (1420–1850 nm and 2000–2200 nm). Similar results have been observed at canopy level for the Sentinel-2 data, particularly in the NIR and SWIR infrared regions (Fig. 5A). Infested trees tend to have a higher reflectance in the VIS
and SWIR regions than healthy trees. Moreover, the results of one-way ANOVA testing shows that the mean reflectance spectra of healthy and green-attacked sample plots were significantly different ($P < 0.05$) for all Sentinel-2 spectral bands at both leaf (simulated spectra) and canopy level. In contrast, for Landsat-8, no significant difference was observed at the canopy level. However, from the simulated leaf spectra of Landsat-8 spectral bands, low significance differences were observed for the red, NIR, SWIR1 and SWIR 2 bands (Table 2).

**Principal component analysis (PCA) and ANOVA test**

Significant differences ($P < 0.05$), were found for all SVIs calculated from Sentinel-2 (except SR-Blue/Red, GDVI and CI) between healthy and infested trees (Fig. 6). Whereas, the SVIs derived from Landsat-8 were only sensitive to changes in water content (water indices) (DSWI, NDII and LWCI), and no significant differences were found for the pigment indices (Fig. 6). Moreover, the results of PCA revealed that the first two components (PC 1 and 2) explained more than 70% of the variance in the samples of SVIs investigated in this study. Figure 7A, B, and C shows the clustering of two sample plots (healthy and infested) within the space of the first two principal components (PC1 and 2). As can be seen from Figure 7 (A), there was a slight crossover between healthy and infested sample plots when all three groups of SVIs from Sentinel-2 data were applied in the model. In contrast, the two sample plots (healthy and infested) exhibited an apparent overlap and mixed scattering for the Landsat-8 SVI groups. Furthermore, for Sentinel-2 SVIs, the majority of healthy plots tended to be on the negative side of PC1, while the infested plots were relatively scattered on the positive side of PC1 (Figure 7B). Notably, no obvious improvement was observed when the PCA model was built independently for each SVI group.

**Mapping bark beetle green attack and validation**

To determine the importance of each of the SVIs when mapping bark beetle green attack, selection probability (SP) values were obtained from the RF algorithm. A total
of 8 out of 24 indices shown in Figure 8 for Landsat-8 indicated a higher SP (>0.5), whereas 17 of 35 SVIs were recorded with an SP value of >0.5 for Sentinel-2 (Fig. 8).

Moreover, higher variation and wider threshold values were observed between these two sample groups (healthy and infested) using the SVIs considered in this study for the Sentinel-2 data than for Landsat-8 (see Fig. 9). Likewise, using raw spectral bands from Sentinel-2 indicates that red, red-edge1-3, NIR and SWIR bands exhibited SP values of >0.5. In contrast, the Landsat-8 red band had an SP value >0.5. From Sentinel-2 data, the indices selected to generate a final infestation map included NDRE-2, NDRE-3, GLI, GNDVI, NDVI, NGRDI, CIG, PGR, DSWI, LWCI, NDWI, SR-SWIR, RDI, red, red-edge1&2, NIR and SWIR1. However, from Landsat-8 data only CIG, GLI, NGRDI, PGR, DSWI, NDWI, DRI and red were selected.

Based on the defined threshold value, infestation maps were generated using the Landsat-8 and Sentinel-2 SVIs data. These maps were then overlaid with ground truth reference data to calculate matched and mismatched pixels. Figure 10 depicts those areas (pixels) correctly matched with ground truth data (aerial photography), as well as pixels that have been falsely identified from SVIs as green-attacked areas. As can be seen from Table 3, the number of correctly matched pixels with reference infestation data (visual interpretation of aerial photography) was higher for Sentinel-2 (67%) than for Landsat-8 (36%) SVI data. Similarly, the number of falsely identified pixels (mismatched pixels) that indicated green attack...
using SVIs was lower for Sentinel-2 (177 pixels) than for Landsat-8 (391 pixels).

**Discussion**

Spectral vegetation indices (SVIs) calculated from Sentinel-2 have a high potential for mapping and detecting changes induced by bark beetle green attack, particularly the red-edge and water-related indices. These changes were only partly detectable by Landsat-8 due to the lower spectral and spatial resolution of the OLI sensor. A greater number of pixels identifying green attack from Sentinel-2 SVIs matched with ground truth data (362 out of 539 pixels), whereas Landsat-8 only matched 221 pixels (out of the 612) with ground truth infestation data. Furthermore, leaf-level comparisons between healthy and green-attacked foliar samples revealed that all leaf traits considered in this study were significantly different \((P < 0.05)\), particularly for chlorophyll and water content. The reduction in chlorophyll and water content in the infested trees caused changes in the foliar spectral measurements at NIR and SWIR wavelengths.

In our study, the stand characteristics between the healthy and infested plots were not significantly different (not shown). This is in line with earlier studies that found stand characteristics do not play a major role in infestation during an outbreak condition. In an epidemic level of bark beetle infestation all trees with different conditions (healthy and stressed) and different stand characteristics (large or small diameter) were under threat of this insect (Wermelinger 2004; Allain et al. 2011; Lausch et al. 2011). However, it is important to note that in our study, DBH with the \(P\)-value of \(<0.06\) was narrowly exceeding the threshold of 0.05, which probably indicate that this variable may play a role in attracting the beetles. Previous findings by Hart et al. (2015) and Six and Skov (2009) indicate that DBH played a major role in the endemic level of bark beetle infestation.

A common observation in this study was that of the 19 pigment-related SVIs (group-1) calculated from Sentinel-2 imagery, eight (i.e. GLI, NDVI, GNDVI, CIG, PGR, NGRDI and both red-edge-related indices (NDRE-2 and 3)) were most important having an SP value of \(<0.5\) (Fig. 8). When only using VNIR bands from Landsat-8, CIG, GLI, NGRDI and PGR have the potential to differentiate between healthy and infested sample plots. The underlying reasons may relate to the existing dissimilarities in spectral and spatial resolution between these two sensors (Mandanici and Bitelli 2016). For example, Landsat-8 has only one spectral band (30 m) in the near-infrared region, while Sentinel-2 has a series of spectral bands (20 m) in the near-infrared region (B5, B6, B7 and B8a). Furthermore, the availability of the three red-edge bands is a unique feature that distinguishes Sentinel-2 from Landsat-8. In our study, the indices developed from the red-edge bands (705–783 nm) of Sentinel-2 (NRED2 and 3) showed the highest sensitivity to bark beetle green attack, and a larger threshold was identified between healthy and infested samples for these two indices (Fig. 8). This result is consistent with results from previous studies (Eitel et al. 2011; Krofcheck et al. 2014; Lottering et al. 2016; Modzelewska et al. 2017), indicating that NDRE calculated from RapidEye and WorldView-2 has the capacity to detect forest stress induced by drought and insect infestation in the early phase.

Interestingly, the spectral indices using the blue band (CI, SR-BLUE/RED and BNDVI) were unable to discriminate trees stressed by bark beetle green attack for both Landsat-8 and Sentinel-2 (Fig. 8). This is due to the reflectance of these sensors in the blue region of the spectrum being insufficient to detect spectral variation caused by bark beetle green attack. This result confirmed earlier
findings by Arellano et al. (2015), who revealed that the blue range indices calculated from Hyperion images were unable to detect forest areas polluted by hydrocarbon in the Amazon rainforest.

On the other hand, the indices calculated from a combination of SWIR and NIR or VIS bands (water-related indices) performed well for both Sentinel-2 and Landsat-8 imagery. For example, Figure 9 suggests that the indices DSWI, NDWI, RDI, SR-SWIR and LWCI successfully differentiated healthy from green-attacked plots using Sentinel-2. Due to the significantly lower water content and higher leaf dry matter content ($P < 0.05$) of infested samples, their spectral reflectance responded more profoundly in the SWIR region at

Figure 7. Cluster plots based on the first two principal components (PCs): (A) cluster plots based on all spectral indices including raw spectral bands, (B) cluster plots based on pigment indices and (C) cluster plots based on water-related indices.
both leaf and canopy level than the reflectance of healthy samples did (Fig. 3) (Wang et al. 2011; Ali et al. 2016). This is because a reduction in leaf water content is responsible for changes in SWIR reflectance (Bowman 1989). This is also demonstrated in Figures 4 and 5A and B, where distinctly higher reflectance in the SWIR region was observed for the infested samples at both leaf and canopy level. This is consistent with the study by Immitzer et al. (2016), which highlighted the importance of the SWIR in Sentinel-2 data for mapping different forest classes. Similar results have been reported for wavelengths in the NIR and SWIR regions, which are sensitive to forest disturbance caused by insect attacks. Foster et al. (2017) found that the SWIR region from hyperspectral measurements collected from ASD FieldSpec Pro was key to the detection of bark beetle (Dendroctonus rufipennis) green attack in Engelmann spruce (Picea engelmannii).

The sensitivity of SWIR bands to variation in leaf water content is due to its reflective nature, allowing it to bounce off objects while remaining invisible to the human eye. This feature of the SWIR region makes variation in leaf water content easily recognizable. Furthermore, it should be noted that not only the SWIR region is sensitive to variation in leaf water content, but that the thermal infrared region may also provide sufficient information in this

Figure 8. Selection probability value of spectral vegetation indices (SVIs) obtained from PLS-DA Random frog algorithm.
regard. Landsat-8 has two spectral bands in the thermal infrared region (10.60–12.51 nm). Further research would, therefore, be useful to assess whether information from the thermal region can provide sufficient information for mapping bark beetle green attack.

Furthermore, the results of PCA analysis revealed that the method using different SVI groups (pigment or water-related indices) did not improve discrimination between healthy and infested sample plots when using Sentinel-2 data (Fig. 7). However, with Landsat-8 data, the performance of the PCA model decreased when SVIs groups were used separately. This reduction in PCA performance was expected due to the spectral variation in infested plots not being efficiently detected by Landsat-8 images (see Fig. 5A and B).

It should also be noted that 177 and 391 pixels indicating green attack based on Sentinel-2 and Landsat-8 data, respectively, were mismatched with the ground truth data (Table 3). To better understand this discrepancy, we studied the infestation data (obtained from visual interpretation of aerial photography) of the previous years, 2014 and 2015, acquired from the Bavarian Forest National Park administration. This allowed us to check whether mismatched pixels indicating green attack corresponded with previous infestations in 2014 and 2015. To accomplish this, a similar process (as explained in Section Ancillary data and accuracy assessment) was applied to 2014 and 2015 infestation data and then overlaid on the mismatched pixels. As can be seen from Table 4, the results revealed that, for Sentinel-2, 12 and 57 pixels (out of the 177 mismatched pixels) matched with the infestation data of the years 2014 and 2015, respectively. While for Landsat-8 data, only 11 and 31 pixels matched with infestation data for the years 2014 and 2015, respectively (Table 4). This highlights the sensitivity of SVIs calculated from Sentinel-2 for detecting canopies that are stressed by bark beetle green attack. It is important to note that all the pixels that were identified as green attack, using both sensors considered in this study, were within 500 meters of the

![Figure 9](image-url). The box plot shows the variation in spectral vegetation indice (SVI) values calculated from Landsat-8 and Sentinel-2 (see the next page), between healthy and infested plots. The black and hollow boxes represent healthy and infested plots, respectively. The red box shows the selected threshold value of each SVI where there is no overlap between the two sample groups (healthy and infested).
Figure 9. Continued.
Figure 10. Map showing the spruce cover under green-attack stress in the Bavarian Forest National Park in July 2016 based on spectral vegetation indices from Landsat-8 and Sentinel-2 selected through the Random frog algorithm.

Table 3. Assessment of the generated maps from Landsat-8 and Sentinel-2 spectral vegetation indices (SVIs) using the reference data obtained from Aerial photography.

<table>
<thead>
<tr>
<th>Identified pixels as green attack from (SVIs)</th>
<th>Reference pixels (Aerial photography)</th>
<th>Pixels correctly matched</th>
<th>Mismatched pixels</th>
<th>Error$^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-8 (612 pixels)</td>
<td>417 (30 m)</td>
<td>221 (36%)</td>
<td>391</td>
<td>64%</td>
</tr>
<tr>
<td>Sentinel-2 (539)</td>
<td>687 (20 m)</td>
<td>362 (67%)</td>
<td>177</td>
<td>33%</td>
</tr>
</tbody>
</table>

$^1$The error was calculated by dividing the total number of correctly matched pixels by the total number of ground truth pixels.
previous year’s infestation, when analyzed using a 500-meter buffer zone around these data (not shown). Hence, we could identify the previous years’ infestations, which were now at the grey-attack stage.

Conclusion

The spectral indices derived from Sentinel-2 data performed well at detecting changes in the leaf biochemical properties (reduction in chlorophyll, increase in leaf dry matter content and decrease in water leaf water content) and their relation to canopy reflectance. The simulated Sentinel-s data also showed good accordance with measured leaf reflectance using an ASD spectrometer as well as a superior response to changes in leaf biochemical properties over the whole wavelength region, as almost all utilized SVIs performed well for detecting bark beetle green attack. Although the total number of pixels matched with ground truth data (362 pixels, 67%) obtained from the Sentinel-2 may not be high enough for operational forestry practice and management purposes, it is a promising technique for alerting to bark beetle green attacks. It can aid the detection of bark beetle infestations in a timely manner over large areas and thus form the basis for accurate and efficient bark beetle monitoring. It is possible that another type of remote sensing data with a higher resolution (such as from an Unmanned Airborne Vehicle) may provide a better result detecting bark beetle green attack. As early detection of infestations is essential for the successful control of an outbreak (Presse and Heurich 2011), further research applying different approaches using Sentinel-2 imagery should be undertaken with principal component analysis (PCA) and partial least square discriminate analysis (PLS-DA) from this study to check the stability and accuracy of the threshold values identified over a period of time.

Acknowledgments

This research received financial support from the EU Erasmus Mundus Salam-2 and was co-founded by Natural Resources Department, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, the Netherlands. The authors thanks the Bavarian Forest National Park staff for approving access to the test site and providing field and camping facilities. The authors extend their appreciation for the great support during the laboratory measurements from the GeoScience Laboratory at ITC Faculty, University of Twente.

References


Analytical Spectral Devices, Inc. 5335 Sterling Drive, Suite A, Boulder, CO 80301-2344, Colorado, USA.


Table 4. Assessment of the mismatched pixels from Landsat-8 and Sentinel-2 spectral vegetation indices (SVIs) using previous years’ (2014 and 2015) infestation data (obtained from Aerial photography).

<table>
<thead>
<tr>
<th>Mismatched pixels as green attack from (SVIs)</th>
<th>Pixels correctly matched with infestation data (2014)</th>
<th>Pixel correctly matched infestation data (2015)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-8 (391 pixels)</td>
<td>11</td>
<td>31</td>
</tr>
<tr>
<td>Sentinel-2 (177 pixels)</td>
<td>12</td>
<td>57</td>
</tr>
</tbody>
</table>


**Supporting Information**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Appendix S1.** Spectral vegetation indices applied to leaf reflectance measurements, Sentinel-2 and Landsat-8 in the study area.