Master Thesis

submitted within the UNIGIS MSc programme
Interfaculty Department of Geoinformatics - Z_GIS
University of Salzburg

Detection of invasive exotic conifers and control measures with Sentinel-2 satellite data

by

Dipl.-Ing. Silv. (Univ.) Sebastian Klinger
105491

A thesis submitted in partial fulfilment of the requirements of the degree of
Master of Science (Geographical Information Science & Systems) – MSc (GIS)

Advisor:

Dr. Christian Neuwirth

Rotorua, 10th October 2022

Acknowledgements

This thesis was partly carried out at Scion (New Zealand Forest Research Institute Ltd.) as part of the programme “Vive la résistance – achieving long-term success in managing wilding conifer invasions” led by Scion (New Zealand Forest Research Institute Ltd.). This research project is an international collaboration between Scion, Lincoln University (NZ), University of Canterbury (NZ), Manaaki Whenua (NZ), Australian National University and the National Centre for Atmospheric Research (USA) to support New Zealand’s National Wilding Conifer Control Programme.

First, I would like to thank my supervisors Dr Christian Neuwirth (University of Salzburg), Dr Grant Pearse (Scion) and Dr Thomas Paul (Scion) for their continuous support, kindness and encouragement during my thesis. I certainly learned a lot and their valuable and critical feedback and discussions helped to improve the quality of the thesis. Thanks to Ben Steer (Scion) for the work on the high-resolution reference datasets and Dr Nicolò Camarretta (Scion) for the helpful discussions on the topic.

Many thanks also to the UNIGIS team and module supervisors for the interesting GIS modules and support during my studies.

Finally, I would like to thank my family and friends for their support, assistance, and patience during my studies. This is very much appreciated!
My sincerest thanks to my wife Vanessa, who supported me through all ups and downs during my studies during a challenging time, and kept my back free.

Science Pledge

I hereby declare that the thesis is entirely the result of my work. I have cited all sources I have used in my thesis; I have always indicated their origin. This thesis was not previously presented to another examination board and has not been published.

Rotorua, 10/10/2022	S. Klinger
Place, Date	Signature

Preface

This thesis is the final part of my studies in Geoinformatics at the University of Salzburg (UNIGIS). It was carried out between February 2022 and October 2022 in New Zealand. It highlights the results of the detection and classification of invasive exotic conifers with Sentinel-2 satellite data using machine learning techniques.

This thesis is submitted as a manuscript-based thesis as it identifies and addresses a research gap. The techniques and results are highly relevant for the ecology and management of invasive species and will help to monitor and manage these. It is intended to publish the manuscript in a high-impact peer-reviewed journal.

The manuscript is written for publication in ‘Forest Ecology and Management’. This journal has a strong focus on forest ecology but aims to link this with forest management. The editors encourage to submit manuscripts with novel ideas or approaches to important challenges in forest ecology and management, and the use of remote sensing technologies has been highlighted in numerous publications. Last, the journal encourages communication between scientists in disparate areas who share a common interest in ecology and forest management, bridging the gap between research workers and forest and land managers. Therefore, this journal is a suitable pathway to publish the results of this research.

This thesis contains two main sections. The first comprehensive section includes the manuscript that will be submitted to the ‘Forest Ecology and Management’ Journal shortly for publication. The short report in section 2 covers the further technical analysis of this research work, which has not been included in more detail in the manuscript. Due to the comprehensive and very detailed manuscript (>11,000 words), this section is relatively short. Most aspects have been covered in the manuscript due to its methodological research question.

List of Contents

A) Manuscript

B) Report

A) Manuscript

Detection of invasive exotic conifers and control measures with Sentinel-2 satellite data

Authors: Sebastian Klinger[*1, 2], Grant Pearse[1], Benjamin S.C. Steer[1], Thomas Paul[1]

1. Scion (New Zealand Forest Research Institute, Ltd.), Titokorangi Drive, Private Bag 3020, Rotorua, 3046, New Zealand

2. PF Olsen Limited, Te Papa Tipu Innovation Park, 99 Sala Street, Rotorua 3010, New Zealand

***** *Corresponding author: sebastian.klinger.nz@gmail.com*

ABSTRACT

Invasive species are a threat to rare ecosystems and can cause adverse effects on indigenous biodiversity, productive land use and landscape experience. In New Zealand, several invasive exotic conifer tree species have been introduced in the last century and have now invaded around 2 million hectares. To manage and control invasive exotic conifers, the New Zealand Government has spent nearly NZ$100 million over the last five years to stop further spread and remove large infestations in many parts of the country. With the ongoing management of invasive conifer infestations on a large-scale, remote sensing techniques offer a promising tool to monitor the development of existing and new infestations and track the success of the national control programme efficiently.

In our study, we tested freely available Sentinel-2 satellite data to detect advanced exotic conifer infestations. We developed pixel-based classifiers to detect areas infested with exotic conifers in Sentinel-2 imagery and compared the accuracy of these classifiers against the accuracy of classifiers developed from high-resolution imagery. An accuracy of 99 % in the validation model could be achieved in the classification of healthy conifer infestations with Sentinel-2 imagery. Conversely, the accuracy of high-resolution classification with WorldView-3 was only 82 % in the validation model. The higher accuracy of Sentinel-2 is attributed to the availability of spectral bands in the SWIR and red-edge spectrum and the higher purity of pixels in the training data. Consequently, classification with Sentinel-2 imagery is a promising tool to detect and monitor advanced infestations of unwanted exotic conifers. The detection of treated exotic conifer infestations was improved during the red-needle stage that occurs shortly after control operations and before needle-fall (grey phase).

Keywords: invasive conifers, remote sensing, Sentinel-2, invasive alien plants, invasive species management, tree detection and mapping, ecology

1. INTRODUCTION

Northern hemisphere conifers, in particular those of the genus *Pinus* have become invasive after their introduction for timber production, soil stabilisation or amenity into the southern hemisphere (Nuñez et al., 2017; Richardson & Rejmánek, 2011). In New Zealand, invasive exotic conifers invade indigenous low-stature ecosystems such as native or semi-native grasslands that often form the basis for food production through grazing. The impact of such tree invasions in New Zealand on the loss of agricultural production, landscape value and risk of habitat loss to indigenous flora and fauna are estimated to be in the billions (Ministry for Primary Industries New Zealand, 2014; Nuñez et al., 2017; Richardson & Rejmánek, 2011). A recent review on ecological research to improve long-term outcomes of invasive conifer management describes different pathways to the desired land use outcomes after control operations (Dickie et al., 2022). Pine species, such as *Pinus contorta (Douglas ex Loudon) (lodgepole pine)*, *P. sylvestris L. (Scots pine)* and *P. nigra J.F. Arnold subsp. nigra (Corsican pine)* are considered invasive weeds in New Zealand. The New Zealand Government runs a large management programme to control 'wilding pines' under the 'New Zealand Wilding Pine Strategy' (Ministry for Primary Industries New Zealand, 2014). Even with substantial efforts to control wilding infestations in the past, wilding pines are still present on around 2 million hectares of New Zealand's unique natural environment (Peltzer, 2018). This area is already larger than New Zealand's total well-managed exotic forest plantation estate with around 1.7 million hectares (New Zealand Forest Owners Association, 2020) and in 2014 the government estimated that the area invaded expands at a rate of 6% per year under the past level of management (Ministry for Primary Industries New Zealand, 2014).

Since the total area affected by wilding pines is large, and infestations are widely distributed, remote sensing approaches to estimate extent and density have been identified as a promising tool to support control programmes by enabling the monitoring of control success and mapping the ongoing spread and intensification of wilding infestations (Nuñez et al., 2017; Visser et al., 2014; Sprague et al., 2019; Dash et al., 2019a). Approaches to describe invasive conifers with different remote sensing data from different platforms like UAV (Dash et al., 2019b), aerial images (Sprague et al., 2019) and laser scanning (Dash et al., 2017) or Landsat satellite imagery (Dash, 2020) have shown great potential to monitor invasion of wilding pines at different spatial scales. To detect large-scale spread of wilding pines or control activities, medium-resolution Landsat satellite imagery has proven to be very suitable (Dash, 2020). Time-series remote sensing data can also inform about the historical patterns and spread through sequential information and detect landscape changes (Dash, 2020). Landsat imagery has also been used to monitor the spread of invasive *Acacia longifolia* in Portugal (de Sa et al., 2017) and several vegetation change detection algorithms based on Landsat time-series data have been developed (Zhu, 2017; Kennedy et al., 2015).

While multi-spectral, very high resolution (VHR) satellite or airborne hyperspectral data have also been used successfully for tree species mapping (Immitzer et al. 2012; Fassnacht et al., 2017; Waser et al., 2014), the use of such approaches operationally across large areas is often limited by high acquisition costs and the difficulty in capturing large areas quickly.

With the launch of Sentinel-2 satellites in 2015 and 2017, freely accessible satellite imagery with a comparatively high spectral, spatial (ten bands with 10-20m and three bands with 60 m) and temporal (3-5 days) resolution (ESA, 2022a) have become available and have proven to be beneficial to many applications in forest environment such as forest mapping (Bolyn et al., 2018), tree species mapping (Grabska et al., 2020; Grabska et al., 2019; Immitzer et al., 2019), land-use and land-cover mapping (Forkuor et al., 2018), forest resource inventory (Malcolm et al., 2021) and small-scale change detection in forests (Grabska et al., 2020). Sentinel-2 data have not been used to date to map areas affected by wilding pines in New Zealand or to evaluate control measures of the National Control Programme. The addition of red-edge bands and the higher spatial resolution of Sentinel-2 imagery results in significant improvements in land-cover mapping compared to the previously used Landsat imagery (Forkuor et al., 2018) and shows promise to detect and monitor exotic conifer infestations in New Zealand.

The aim of this study was to test freely available Sentinel-2 satellite imagery to detect wilding pine infestations of various densities and sizes in a heterogeneous landscape, their development over time and to test whether Sentinel-2 data are suitable for tracking land cover changes associated with historical invasive conifer spread and the control of conifer infestations in New Zealand.
In particular, the objectives of this research were (A) to develop a pixel-based classifier to detect wilding pines in Sentinel-2 imagery using spectral bands and common vegetation indices, (B) to evaluate how accurate this classifier was at detecting invasive exotic conifers in natural and semi-natural ecosystems compared to a classifier developed from using high-resolution imagery from the WorldView-3 satellite and (C) test if this classifier can be used to detect successful control activities in previously invaded areas by comparing pre and post control Sentinel-2 images.

2. MATERIALS & METHODS

2.1. Study area

The study site ranged from 540 to 1530 m in elevation, with slopes up to 41° (Table 1). Mean annual temperature ranged from 4.2 – 9.5°C, and mean annual rainfall between 800 and 1000 mm (Table 1); Soils included mostly acidic allophanic brown soils. The vegetation is predominantly low-producing or depleted grasslands, dominated by exotic grasses such as *Agrostis* spp., *Anthoxanthum odoratum* with some native grass species also present (*Poa colensoi, Festuca novae-zelandiae*). Tall tussock grasslands (*Chionochloa* spp.), smaller shrubs such as matagouri (*Discaria tomentosa)*, manuka (*Leptospermum scoparium*) and kanuka (*Kunzea robusta*), indigenous mountain beech forests (*Fuscospora cliffortioides*), exotic forests, shelterbelts and invasive exotic pines are also part of the vegetation mosaic of the study area. Unvegetated areas such as open soil, gravel and rocks are also part of this typical mountainous landscape on the eastern side of the South Island Great Divide (Wardle, 1991).

Table 1: Site parameters related to growing conditions, including altitude (Landcare Research NZ Ltd, 2020a), slope (Landcare Research NZ Ltd, 2020b), mean annual temperature (Landcare Research NZ Ltd, 2020c), average annual rainfall (Ministry for the Environment and Statistics NZ, 2017), and soil class (Landcare Research NZ Ltd, 2018; Hewitt, 1998)

Region	Latitude	Longitude	Altitude range [m]	Slope [°]	Mean annual temperature [°C]	Average annual rainfall 1972-2016 [mm]	Main Soil class
Canterbury, New Zealand	-43.1667°	171.7445°	540-1530	0 - 41	4.2 – 9.5	800 - 1000	Acidic Allophanic Brown Soils

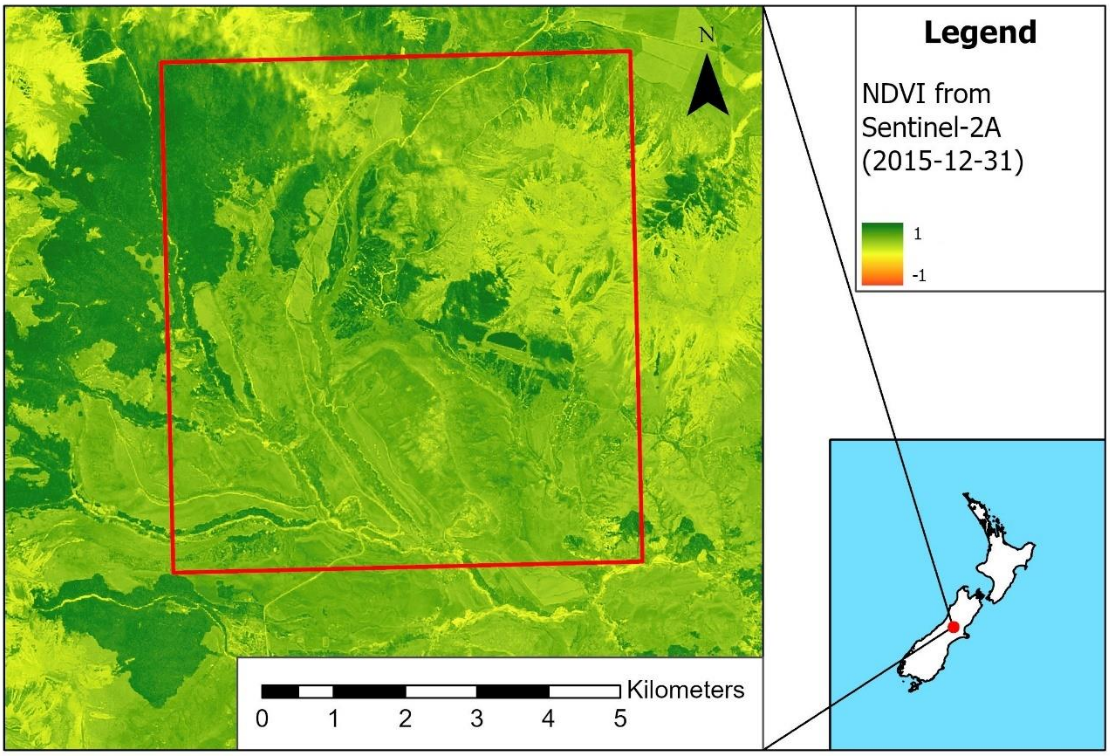

Figure 1: Map of the study area.

2.2. Methods

In this study, we used four different satellite imagery datasets. One pair of images was taken in summer 2015/2016 and the other pair in summer 2022. Each pair consists of one high-resolution satellite image used as a reference for ground truth and training areas as well as evaluation and one Sentinel-2 image. The exact recording dates are shown in Table 2. Figure 2 provides an overview and a workflow diagram of the methods used to process the imagery in this study.

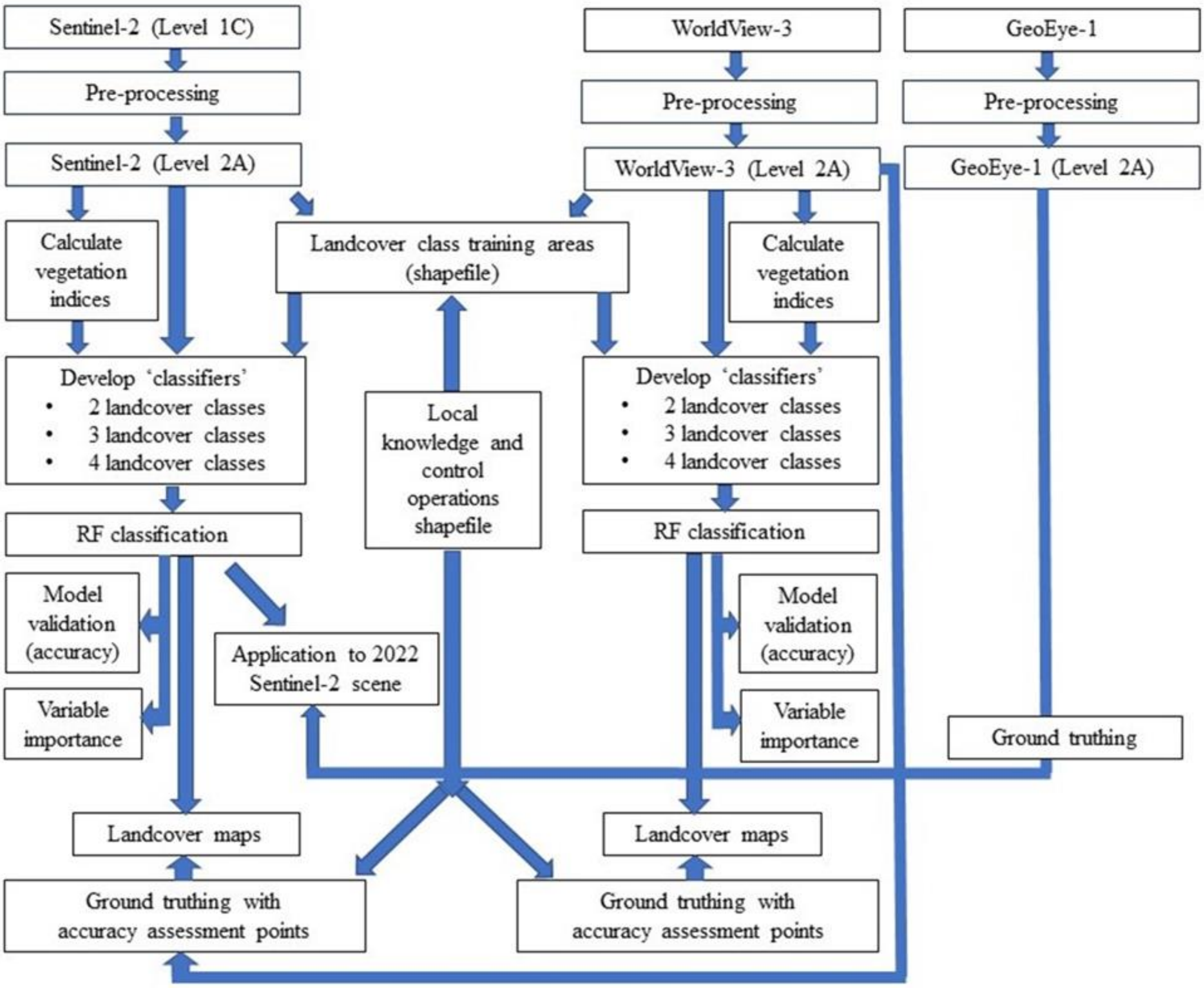

Figure 2: Workflow diagram of methods

Classifiers were developed from the 2015 Sentinel-2 and the WorldView-3 satellite data to compare the Sentinel-2-derived classification against a high-resolution classification. 'Healthy pines' as well as already controlled 'unhealthy pines' are both present in these early scenes. The second pair of images (summer 2022) was recorded after all pines in the scene had been controlled with the purpose of elimination and only a few 'unhealthy' dying or dead pines remained.

Control operations of wilding pines span from manual control such as pulling and cutting, ground-based chemical application methods (knapsack or pasting) to aerial spot spraying of individual trees and larger scale boom spray operations of dense stand-like infestations using helicopters (Ministry for Primary Industries New Zealand, 2014). When herbicide-based control methods are applied, conifer needles initially turn red ("red phase"). With time (~1 year), the dead trees cast their needles and the tree skeleton remains ("grey phase"). Therefore, controlled trees will first appear green-red, red, and finally grey in an RGB image.

The Sentinel-2 classifiers developed from the 2015 scene were later applied to the 2022 Sentinel-2 scene to estimate the extent and success of the control operations of wilding pines which occurred after the first scene was taken. The second high-resolution dataset (GeoEye-1) was used to assess the quality of the predicted classification map for the Sentinel-2 imagery from 2022 with the classifier developed with the Sentinel-2 2015 imagery.

Table 2: Imagery used in the study

Description	Satellite	Date
Only some pine areas have been controlled	Sentinel-2 WorldView-3	31/12/2015 21/01/2016
Post large-scale control operations	Sentinel-2 GeoEye-1	13/01/2022 27/02/2022

2.2.1. Sentinel-2 data and environmental variables

The freely available Sentinel-2 satellite data used in our study (Level-2A product for 2022; Level-1C product for 2015; tile number T59GNN;) were downloaded from the Copernicus Open Access Hub (https://scihub.copernicus.eu/). Since Level-2A data are only available from January 2018 outside Europe (ESA, 2022a), Level-1C products were retrieved from the Copernicus Hub for the 2015 imagery and processed to Level-2A using the Sen2Cor package. Version 2.5.5 of Sen2Cor was used, as this is recommended for older Sentinel-2, Level-1C data products. This pre-processing includes atmospheric, terrain and cirrus correction of Top-Of-Atmosphere Level 1C input data to an orthoimage Bottom-Of-Atmosphere (BOA) corrected reflectance product (Level-2A) (ESA, 2022a).

We used manually selected, cloud-free imagery to avoid the impact of clouds and cloud shadows in the area of interest (AOI). In the first scene from 2015, only a small area of pines had been subject to control activities in the AOI, the second image was recorded in 2022 after a larger area had been subject to control operations. Their exact acquisition dates and the corresponding high-resolution imagery used for reference and evaluation are summarized in Table 2. We selected all Sentinel-2 bands with 10 m and 20 m spatial resolution (Table 3). Resampling of 20 m bands to 10 m was performed in R version 4.2.0 (R Core Team, 2022) using the raster package (Hijmans, 2022) and the bilinear interpolation method.

Vegetation indices calculated from combinations of bands have been successfully used to extract meaningful information about land cover and many different vegetation indices have been developed (Bannari et al., 1995; Schindler et al., 2021). In this study, we tested nine vegetation indices summarised in Table 4. All calculations were performed in R version 4.2.0 (R Core Team, 2022) using the 'raster' (Hijmans, 2022), 'RSToolbox (Leutner et al., 2019)', 'rgdal' (Bivand et al., 2015), 'sf' (Pebesma, 2018), 'random forest' (Liaw & Wiener, 2002), 'caret' (Kuhn et al., 2020) packages and visualised using 'ggplot2' (Wickham, 2016) and ArcGIS Pro (Esri, 2022).

Table 3: Sentinel-2 spectral bands used in this study. Indicative average central wavelengths and bandwidths for Sentinel-2 A & B bands (ESA, 2022a)

Band	Short Name	Central Wavelength	Bandwidth	Spatial Resolution
02	Blue	492 nm	66 nm	10 m
03	Green	559 nm	36 nm	10 m
04	Red	665 nm	31 nm	10 m
05	Red-edge 1	704 nm	15 nm	20 m
06	Red-edge 2	740 nm	15 nm	20 m
07	Red-edge 3	782 nm	15 nm	20 m
08	Near Infrared (NIR) wide	833 nm	106 nm	10 m
8A	Near Infrared (NIR) narrow	864 nm	21 nm	20 m
11	Shortwave Infrared (SWIR) 1	1612 nm	92 nm	20 m
12	Shortwave Infrared (SWIR) 2	2194 nm	180 nm	20 m

Table 4: Spectral/vegetation indices used in this study

Index	Sentinel-2 band combinations	Description	Reference
EVI	2.5 * (B8A - B04) / (B8A + 6 * B04 - 7.5 * B02 + 1)	Enhanced Vegetation Index	Huete et al. (2002); Nagler et al. (2005)
GI	B03 / B04	Greenness Index	Smith et al. (1995)
MSI	B11 / B08	Moisture Stress Index	Serrano et al. (2000)
NBR	(B08 - B12) / (B08 + B12)	Normalised Burn Ratio	Key and Benson (1999); Roy et al. (2006)
NDVI	(B08 - B04) / (B08 + B04)	Normalised Difference Vegetation Index	Rouse et al. (1974); Tucker (1979)
RESRI; SR 735/710	B06 / B05	Red-Edge Single Ratio Index; Simple Ratio 735/710	Zarco-Tejada et al. (2001); Gitelson et al. (1999)
SIPI 3	(B08 - B02) / (B08 - B04)	Structure Intensive Pigment Index 3	Blackburn (1998)
TGI	0.5 * (190 * (B4 - B3) - 120 * (B4 - B2))	Triangular Greenness Index	Hunt et al. (2013)
VARIgreen	(B03 - B04) / (B03 + B04 - B02)	Visible Atmospherically Resistant Index Green	Gitelson et al. (2002); Gitelson et al. (2001); Gitelson et al. (2003)

Normalised Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and Moisture Stress Index (MSI) are frequently used in land cover classification (Guyot & Baret, 1988; Huete et al., 2002; Nagler et al., 2005; Rouse et al., 1974). The Normalised Burn Ratio (NBR) has been used frequently to assess forest disturbance severity, especially fire detection (e.g. Roy et al., 2006; Veraverbeke et al., 2010), but was also used recently for the detection of invasive conifers with Landsat imagery by Dash (2020). The simple Greenness Index (GI) was used as visualisation for the discolouration of green plants and as an approximation of the visual observation (Moya & Flexas, 2012; Scholten et al., 2019). Two more visible band indices were used in this study: Visible Atmospherically Resistant Index Green (VARIgreen) (Gitelson et al., 2002; Gitelson et al., 2001; Gitelson et al., 2003) and the Triangular Greenness Index (TGI) (Hunt et al., 2013) which proved to be highly suitable to detect leaf chlorophyll content (Hunt et al., 2013). The Structure Intensive Pigment Index 3 (SIPI 3) is suitable for the analysis of vegetation with variable canopy structure (Peñuelas et al., 1995; Fusaro, 2019). It has been used to describe the stress level of vegetation and related vegetation health, and for identifying crop diseases (Tong & He, 2017). The Red-Edge Single Ratio Index (RESRI) or Simple Ratio 735/710 makes use of the red-edge bands and is described as a precise indicator of chlorophyll content of plant leaves (Gitelson et al., 1999).

2.2.2. High-resolution reference data

A high-resolution evaluation dataset (WorldView-3) as well as local knowledge was used to test the quality of the classifier and to make sure the training and validation dataset are accurate. We also developed a high-resolution classifier derived from all eight multispectral bands and four calculated vegetation indices (NDVI, GI, SIPI 3, TGI) to compare the performance against the Sentinel-2 classifier.

Table 5: WorldView-3 spectral bands in Visible Near Infrared (VNIR) used in this study (ESA, 2022b).

Band	Short Name	Spectral band	Spatial Res. (max.)
1	Coastal Blue	400 - 450 nm	1.24 m
2	Blue	450 - 510 nm	1.24 m
3	Green	510 - 580 nm	1.24 m
4	Yellow	585 - 625 nm	1.24 m
5	Red	630 - 690 nm	1.24 m
6	Red-edge	705 - 745 nm	1.24 m
7	Near Infrared (NIR) wide	770 - 895 nm	1.24 m
8	Near Infrared (NIR) narrow	860 - 1040 nm	1.24 m

A second high-resolution image (GeoEye-1) was used to evaluate the predicted classification of the Sentinel-2 imagery with the classifier developed using the 2015 Sentinel-2 imagery. GeoEye-1 was the only available and suitable image with a small temporal gap to the Sentinel-2 imagery.

The WorldView-3 multispectral image used in this study was acquired on 21 January 2016, with 1.24 m spatial resolution (Table 5). The image was acquired with a mean off-nadir view angle of 17.7° and mean solar azimuth and elevation angle of 163.9° and 56.3°, respectively. The WorldView-3 satellite sensor provides eight multispectral bands covering four commonly used spectral band (blue, green, red, and near-infrared wide) and four less common bands (coastal blue, yellow, red-edge, and near-infrared narrow). Atmospheric correction of the data was performed by the provider (DigitalGlobe, now Maxar) to obtain surface reflectance. This was achieved using a physically-based normalization of the digital numbers. The multispectral product was then pan-sharpened using the Gram-Schmidt algorithm, resulting in a higher spatial resolution of 0.30 m, increasing the efficacy of visually identifying wilding pines.

2.2.3. Classification

Three classification strategies were tested in this work. First, the AOI was classified into two different landcover classes: 'healthy pines' (green in RGB image) and 'other'. The 'other' class included all remaining landcover types present in the scene, such as native forests, unhealthy pines, grassland, shrublands, roads or soil/rocks or water.
In the second classification strategy, the AOI was classified into three different landcover classes: 'healthy pines', 'unhealthy pines (red and brown-grey in an RGB image)' and 'other'. Examples showing the different classes and the stages of 'unhealthy pines' are shown in Figure 3.
Similar to the first classification approach, the 'other' class in the second strategy included all other land covers present in the scene such as native forests, grassland, shrublands, roads or soil/rocks or water. The third classification strategy consisted of four landcover classes: 'healthy pines', 'unhealthy pines', 'native forest' and 'other'.

Training data for use in the supervised machine learning algorithm were delineated based on visual analysis of high-resolution imagery combined with local site knowledge and records and maps of control operations which have been conducted over the years. These input data were split randomly into training and validation subsets in a ratio of 70:30. The training and validation data consisted of up to 106 polygons, split into 27 'healthy pines', 23 'unhealthy pines', 14 'native forests' and 42 'other' polygons respectively.
The same classification strategies were deployed to develop the high-resolution classifier from WorldView-3 imagery. 31 training polygons were slightly shifted spatially to reflect the distortion of the imagery compared to the Sentinel-2 imagery and to make sure the same area of landcover class was used for each training area.

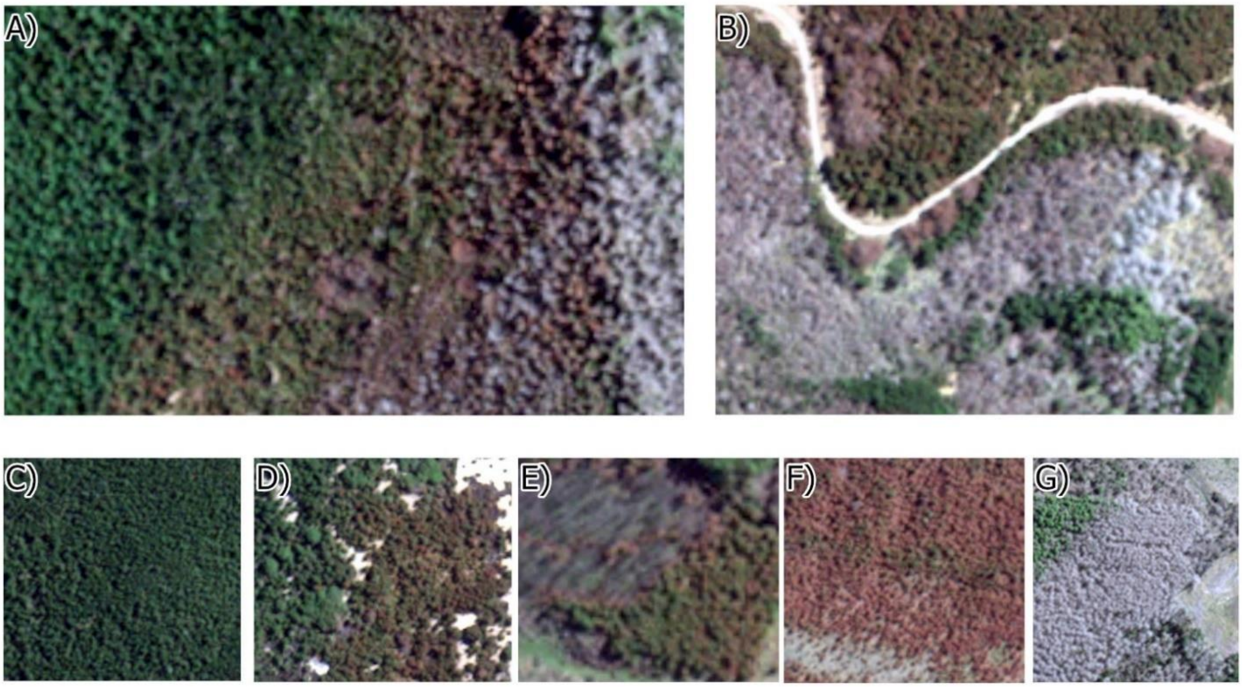

Figure 3: Health classes of wilding pines as defined in this paper. A) shows a transition from 'healthy pines' (left) through the different stages of 'unhealthy pines' (red-green, red/brown, grey); B) points out the difficulty for classification with an intermix of different health classes; C) shows the optimal 'healthy pine' class whereas D)-G) are all part of the 'unhealthy pines' class at different times post control operations

2.2.4. Random Forest model

The supervised machine learning algorithm Random Forests (RF) was used in this study to perform classification of ‘healthy pines’, ‘unhealthy pines’, ‘native forests’ and ‘other land cover’ using the training data to predict categories for the AOI. This method was chosen, because there are many successful use cases in remote sensing using the RF algorithm (Pal, 2005; Forkuor et al., 2017; Grabska et al., 2020; Immitzer et al., 2012; Immitzer et al., 2019; Pullanagari et al., 2018; Einzmann et al., 2021). While single decision trees have the tendency for overfitting, the RF algorithm combines a long list of single decision trees which is more robust to overfitting. It uses a random subset of variables at each decision node and the number of decision trees and variables are the main model input parameters (Breiman, 2001; Liaw & Wiener, 2002). It also performs bootstrap aggregation (bagging) and identifies the relative importance of features during classification (Breiman, 2001). Other recent studies demonstrated that the RF classifier was robust and produced consistently good results compared to other classification algorithms (Persson et al., 2018; Belgiu & Drăgut, 2016) and is even robust and effective with unbalanced data (Breiman, 2001; Pal, 2005; Hastie et al., 2009).

Satellite spectral band values and calculated vegetation indices from the pixels within the training areas were used as predictors in the RF model. We used the ‘superClass’ function from the RStoolbox package (Leutner et al., 2019) in R and tested different settings for cross-validation (*k*-fold) and a number of tuning parameter combinations (tuneLength). In this paper, we present the settings which achieved the highest accuracy of the models, a 5-fold cross-validation and a tune length of three. Random numbers were created (set.seed(5)) for the simulations to ensure all results are reproducible.

To assess the model accuracy, confusion matrices, overall producer’s and user’s accuracy as well as kappa values were computed (Appendix, Tables 7-12).

2.2.5. Variable importance assessment

Variable importance assessment is an important step for high dimensional data and a key step in comparing Sentinel-2 with WorldView-3 data. The RF model supports two methods, Mean Decrease Gini (MDG) and Mean Decrease Accuracy (MDA), to estimate the importance of each predictor variable in the model (Dash et al., 2017). We selected the MDG method as one of the most popular variable importance assessment methods which provides more robust results compared to the MDA (Calle and Urrea, 2011). Out of bag (OOB) sample data are used to calculate variable importance in the RF model (Mellor et al., 2013; Cutler et al. 2007). MDG describes how important each variable is and is calculated by summing all of the decreases in Gini impurity at each tree node split, normalised by the number of trees (Criminisi et al., 2012; Mellor et al., 2013; Dash et al., 2017). Higher values of MDG scores show higher importance of the variable in the model. Variable importance assessment was performed using the ‘varImp’ function of the caret package (Kuhn et al., 2020) in R.

2.2.6. Ground truthing

The validation of the classification (ground truthing) was performed using accuracy assessment points and a confusion matrix. A total of 130-150 accuracy assessment points outside the training and validation polygons were randomly created in ArcGIS Pro using the “equalised stratified random” method. Each point was assessed for each classification approach (2-4 classes) and satellite data (Sentinel-2 and WorldView-3); "Ground truthing", i.e., determining the correct class for each point, was done using the high-resolution WorldView-3 and GeoEye-1 pan-sharpened imagery as a reference, local knowledge of the vegetation and site and records and maps of control operations. A confusion matrix was calculated from both the classified and the reference points which can be found in the Appendix (Tables 13-18).

3. RESULTS

3.1. Classification results

The classification maps are shown in Figures 4-6. Classification of the Sentinel-2 imagery (Figure 4) gives a very good agreement with the reality, showing a large native forest (Figure 4B) in the northwest and 'healthy' and 'unhealthy pines' predominantly in the centre.

The high-resolution classification of the WorldView-3 imagery (Figure 5) provided a fine delineation of vegetation (e.g., shelterbelts) and other landcover classes. However, the classification and separation of vegetation classes ('healthy pines', 'unhealthy pines' and 'native forest') is not as accurate as in the Sentinel-2 classification. The native forest in the northwest for instance is misclassified as 'healthy' or 'unhealthy pines' in several cases.

The predicted classification of the landcover of the 2022 Sentinel-2 imagery (Figure 6) with the classifier developed from the 2015 Sentinel-2 imagery delineates 'unhealthy pines' well, but also confused 'native forests' with 'healthy pines' in several cases, especially in the north-western part of the map. The eradication of 'healthy pines' through control operations is not reflected in the predicted map.

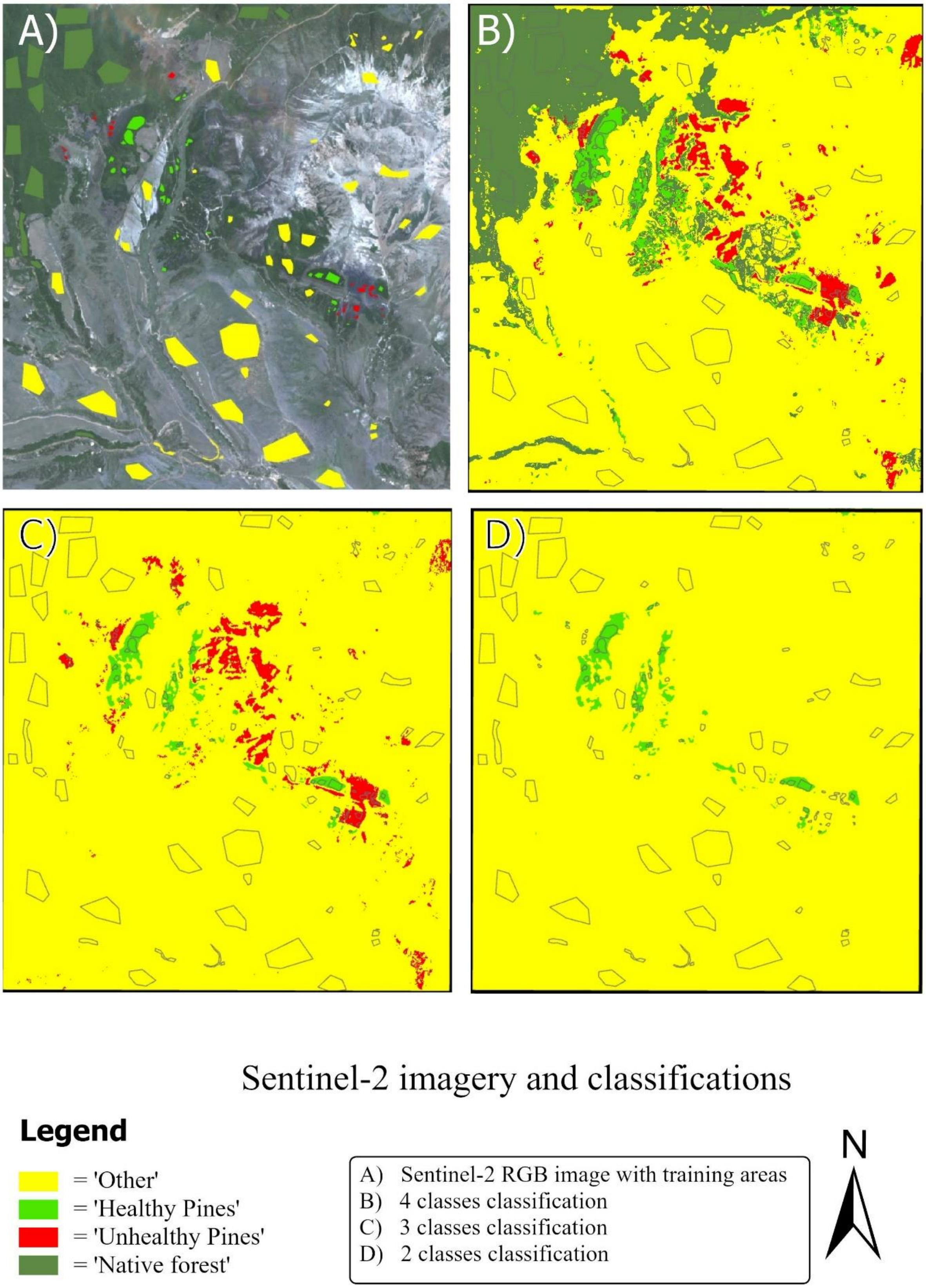

Figure 4: Map classification results of Sentinel-2 imagery

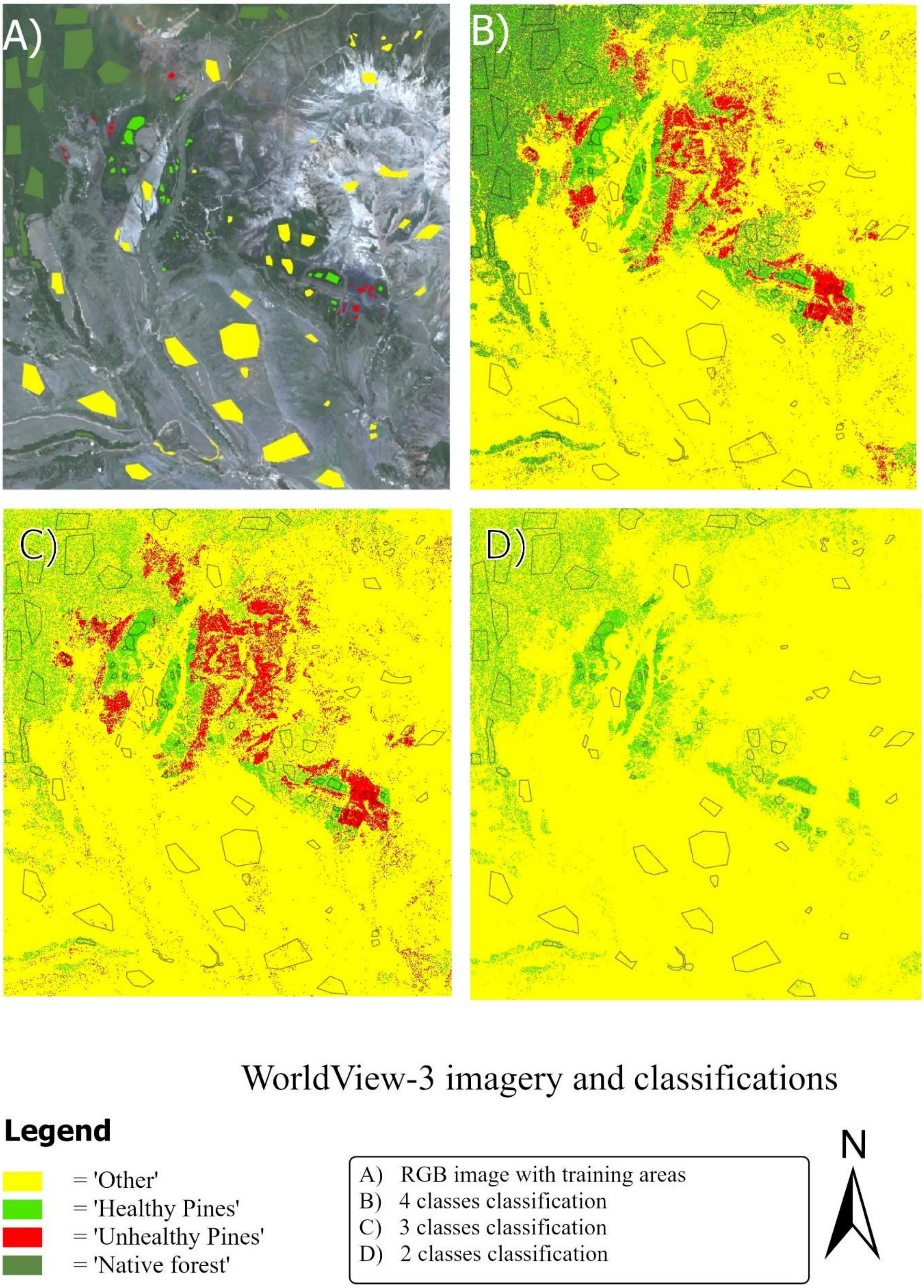

Figure 5: Map classification results of WorldView-3 imagery

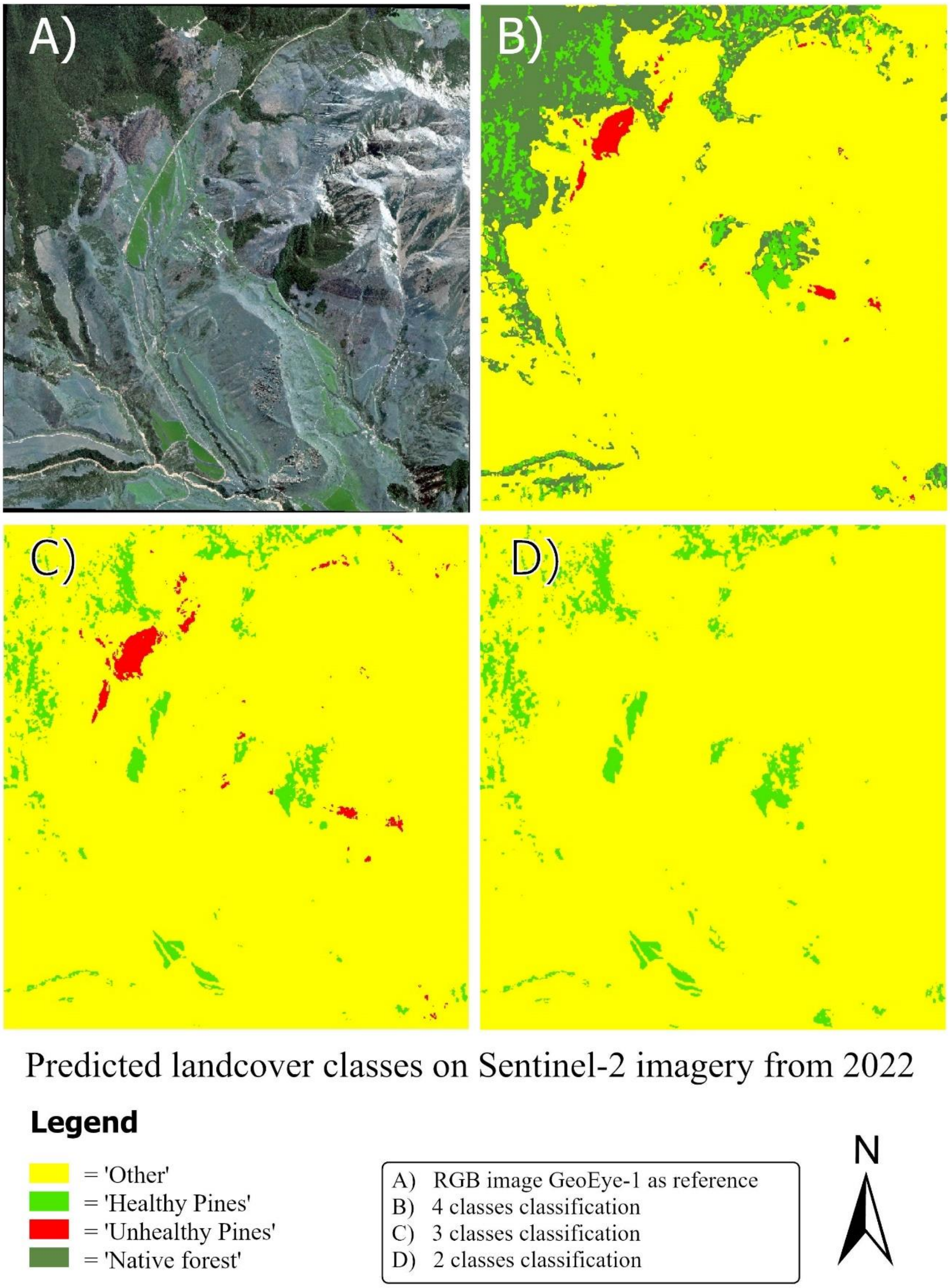

Figure 6: Classified 2022 Sentinel-2 imagery scenes developed with the 2015 classifier (predicted classification)

3.2. Random Forest model performance

The performance of the model for Sentinel-2 imagery is very high with an accuracy of 0.99 and 0.96 and kappa values between 0.96 and 0.92 (Table 6). In contrast, accuracy and kappa values of WorldView-3 image classification models were lower with accuracy between 0.80 and 0.82 and kappa between 0.62 and 0.72 respectively.

Table 6: Validation summary of different classification approaches with Sentinel-2 and WorldView-3; sensitivity classes: 1= 'other'; 2 = 'healthy pines'; 3 = 'unhealthy pines'; 4 = 'native forest'; detailed confusion matrices are provided in the appendix (Tables 7-12)

Satellite	Classification approach	Accuracy	Kappa	Sensitivity	Sensitivity (classes)				Specificity	Specificity (classes)			
					1	2	3	4		1	2	3	4
Sentinel-2	2 classes	0.99	0.96	1					0.93				
	3 classes	0.99	0.94		0.99	0.92	1.00			0.95	1.00	0.99	
	4 classes	0.96	0.92		0.95	0.94	0.99	0.99		0.99	1.00	0.99	0.96
WorldView-3	2 classes	0.82	0.62	0.83					0.81				
	3 classes	0.80	0.67		0.78	0.81	0.84			0.83	0.85	0.98	
	4 classes	0.80	0.72		0.85	0.86	0.85	0.47		0.90	0.88	0.99	0.95

3.3. Variable importance assessment

Sentinel-2

The most important variables for the classifications of the Sentinel-2 imagery are RESRI, NBR and band 11 (SWIR 1). NDVI and B12-band (SWIR 2) were also important in all three approaches (Figure 7). All these vegetation indices and bands are in the red wavelength spectrum. While RESRI uses red-edge bands, the NBR as well as band 11 and band 12 use the short-wave infrared wavelength. Near-infrared bands ranked low in importance with band 8 always around the least important and band 8A mostly among the lowest important bands.

'Green' vegetation indices like TGI, VARIGREEN and GI were of higher importance in the discrimination of the 3-class classification approach where 'healthy pines' were separated against 'unhealthy pines' and the rest of the landcover classes.

WorldView-3

In the absence of any shortwave infrared bands and indices, GI, NDVI and TGI were very important in the classification approaches of the high-resolution imagery (Figure 8). Red and blue bands were the least important in most approaches.

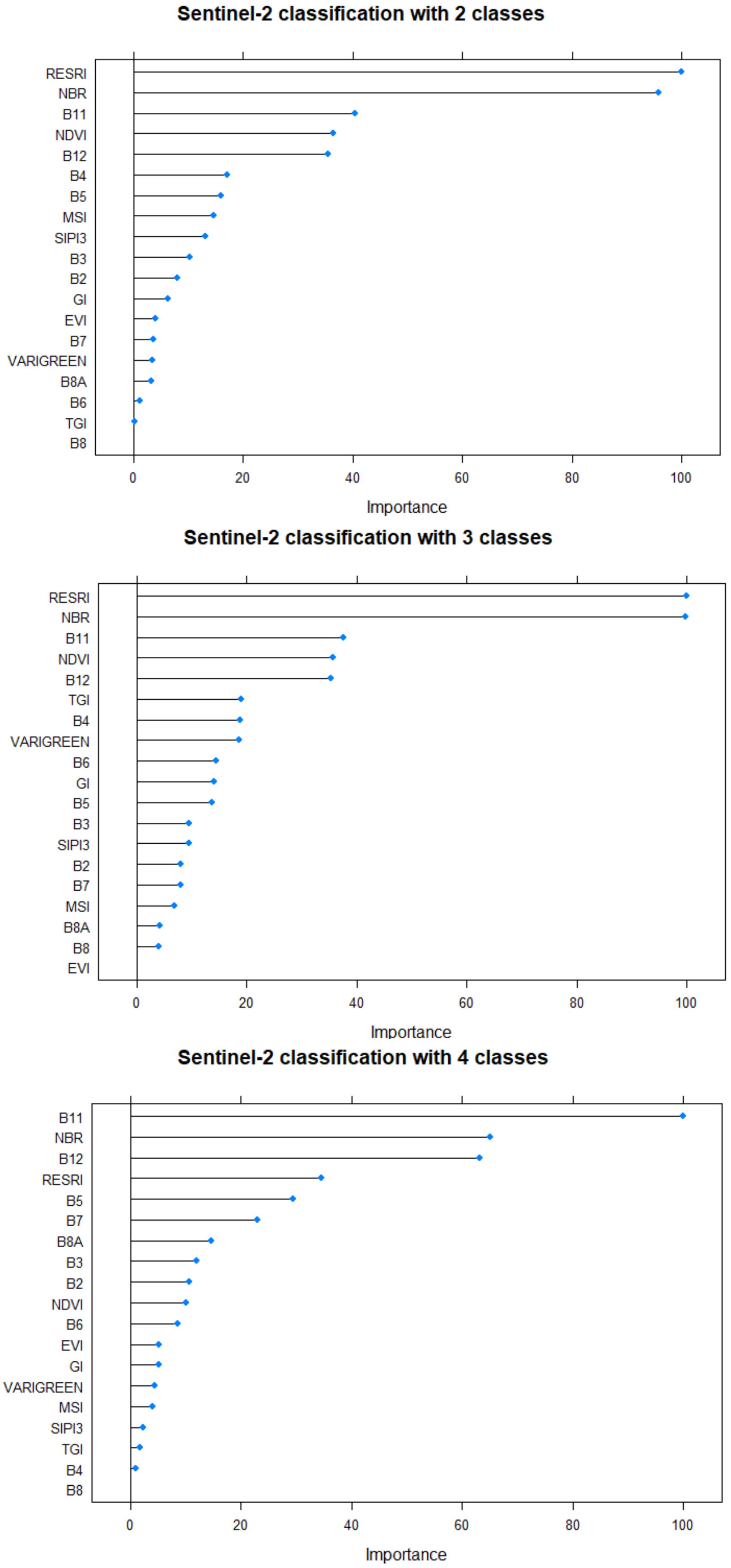

Figure 7: Variable importance of Sentinel-2 classification approaches; the variable importance automatically scales the importance scores to be between 0 and 100 (Mean Decrease Gini) (Kuhn et al., 2020).

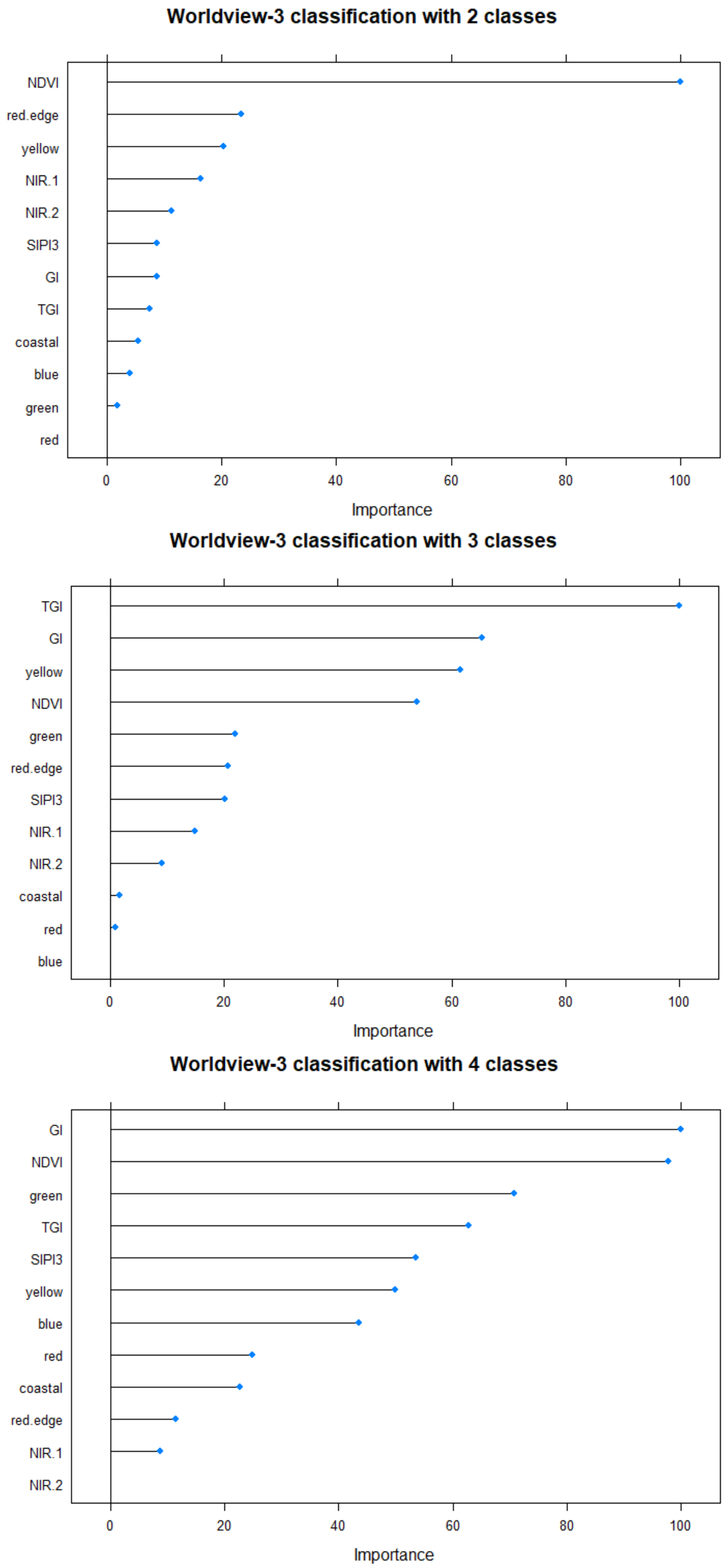

Figure 8: Variable importance of high-resolution classification approach (WorldView-3)

3.4. Ground truthing

The results of the ground truthing in the format of a confusion matrix can be found in the Appendix. Overall user's accuracy (U_Accuracy column) and producer's accuracy (P_Accuracy column) for the 2-class classification of Sentinel-2 was very high with an accuracy of 96% and 91% kappa respectively. Overall user's accuracy for the 3-class Sentinel-2 classification is at 76 % and overall kappa at 63 %, which was the least accurate classification of the Sentinel-2 classification approaches. The 4-class classification resulted in an accuracy of 82% and a kappa value of 76%. The most confused classes in the Sentinel-2 classification were the 'unhealthy pines' class.

High-resolution classification achieved significantly lower accuracy and kappa values. The values for the 2-class classification resulted in an accuracy and kappa value of 68 % and 33 % respectively. Overall accuracy for the 3-class classification is at 63 % and overall kappa at 45 %, which was the least accurate classification of the high-resolution classification approaches. The 4-class classification resulted in an accuracy of 71 % and a kappa value of 61 %. 'Healthy pines' and 'native forest' were the most confused classes in the high-resolution classification.

4. DISCUSSION

4.1. Assessment of the results of Sentinel-2 imagery classification

The classification of wilding pines from Sentinel-2 imagery worked well with comparably high accuracy between 96-99% in the validation of the model. Ground truthing accuracy were between 76% (3-classes classification) and 96% (2-classes) accuracy respectively, which confirms the high accuracy results of the model. Also, from a visual assessment, the recognition of healthy pines worked very well, covering all bigger areas where wilding pines occurred. The separation between native forests and pines achieved reliable results too, only very young trees were mixed up on rare occasions. These results suggest that coniferous wilding pines and native tree species can be well separated with Sentinel-2 data.

The large area of native forests *(fuscospora cliffortiodes*) in the north-west of the imagery was successfully classified as native forests in the 4-classes classification approach, and in the case of the 2- and 3-classes classification correctly classified as 'other' landcover. However, some areas of native forests were confused with an early stage of 'unhealthy pines' when there was a mix of green healthy pines with some interspersed red 'unhealthy pines'. This is potentially because dead trees were present in native forests as well.
The lower accuracy of 76% in the 3-class classification approach was the result of the difficult classification of 'unhealthy pines'. This class is more problematic than the other classes due to the variability and transition phases of 'unhealthy pines', which also makes it even hard for an independent assessor to classify these areas, i.e., check the classification for a 10 m x 10 m pixel. Different components like tree age, local site conditions, shadowing effects, and crown health can affect stand and species spectral properties (Grabska et al., 2021) and spectral variability can be large even within species (Leckie, 2005). Especially the classification of unhealthy trees has been found to be more difficult caused by the variety of stress symptoms possible (Leckie, 2005).

Figure 9 shows the intermixed areas of 'healthy pines', unhealthy reddish and grey pines (which were still counted as 'unhealthy pines') in very close proximity to gaps of grassland or bare soil. Furthermore, trees in some of the controlled areas were poisoned several years ago which resulted in a sparse vegetation cover dominated by other non-pine vegetation or open soil. These areas were occasionally still classified as 'unhealthy pines'. Additionally, a few steep, rocky areas with shadows and pine cover combined resulted in a misclassification as 'unhealthy pines'. These errors resulted in an area overestimation of 'unhealthy pines' and subsequently a greater area for this landcover class. However, the areas that were controlled only recently (which appear red or brown with clearly recognisable pine shapes in a high-resolution RGB or false colour image) were identified very well and no major area was missed. The definition of 'unhealthy pines' class could have been restricted to only green-red, red, and brown pines as unhealthy and the grey and presumably dead pines that were controlled a longer time ago could have been excluded from this class which would have achieved much higher accuracy and less confusion with other dead vegetation.

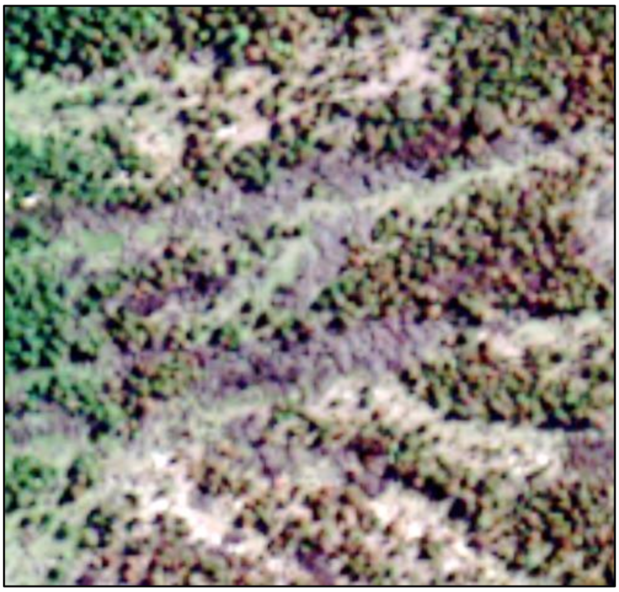

Figure 9: High-resolution WorldView-3 RGB-imagery: Different pine "health classes" (green (healthy) – red-green (unhealthy)– red (unhealthy)– brown (unhealthy) – grey (unhealthy)) in close proximity to each other is challenging to classify in a 10m x 10m Sentinel-2 image.

Variable importance

Key for the overall very successful classification of healthy pines in Sentinel-2 imagery were the infra-red (IR) bands. Especially the shortwave IR bands i.e., B11 and B12, as well as the red-edge bands B5 and B6, were responsible for the high accuracy and good classification result as indicated by the consistently high ranking in the variable importance assessment. These four bands were also used to calculate the most important vegetation indices NBR and RESRI in this study (see variable importance assessment). These results are in accordance with previous studies, which have also highlighted the importance of these bands (SWIR and red-edge) for species classification (Nelson, 2017; Immitzer et al., 2019; Bolyn et al., 2018; Grabska et al., 2019; Puletti et al., 2018; Persson et al., 2018; Immitzer et al., 2016). For broadleaf species classification, Immitzer et al. (2019) describe Sentinel-2 SWIR bands B12 and B11 followed by the red-edge band 1 (B5) as most important and the NIR band as most useful to separate between coniferous and broadleaf trees. Red-edge 2 (B6), SWIR 1 (B11) and SWIR 2 (B12) bands were most important in Persson et al. (2018). The variable importance of their study also reflects the multitemporal approach with scenes analysed over different seasons with the most important variables (SWIR and red-edge) coming from different months (May, October, April).

Differentiation of species and acquisition date

The Sentinel-2 scene was acquired at the end of December which is comparable to an acquisition at the end of June in the northern hemisphere. In the northern hemisphere, single scenes captured in May have been found to be more suitable compared to those from late summer and autumn (Immitzer et al., 2019; Persson et al., 2018; Lim et al., 2020). For the southern hemisphere, this would mean November as the optimal acquisition month which covers the time of spring capturing most phenological differences between the species. However, the alpine environment in this study area results in a later growth flush for both pines and native species. The focus of this study was constrained to separating pines from native forest rather than fine-grained species classification and the methods demonstrated here achieved high accuracy for this purpose.

Capturing changes in tree phenology (i.e., flowering, leaf-onset, and senescence) by satellite scenes is another very successful method. Several studies analysed a multitemporal approach, where the same area is captured over one or more years in different phenological stages which provides differing spectral reflectance. Results showed that accuracy could be increased by analysing satellite images in different seasons (Bolyn et al., 2018; Grabska et al., 2019; Puletti et al., 2018; Persson et al., 2018; Nelson, 2017; Wessel et al., 2018; Hościło & Lewandowska, 2019; Pasquarella et al., 2018; Elatawneh et al., 2013; Hill et al., 2010; Immitzer et al., 2019). However, the seasonal effect might be weaker in this case as native trees in this area are evergreen.

The difference in pixel resolution between Sentinel-2 and Worldview-3 also means a difference in purity of the pixels when the same training areas are used. While the Sentinel-2 training areas consist of ‘pure wilding’ pixels of 10 m by 10 m, the resulting training areas in a high-resolution WorldView-3 image (1.24 m by 1.24 m resolution) also consist of mixed pixels as also spaces between single wilding pines are included in the training areas. This occurrence of isolated pixels with high local spatial heterogeneity between neighbouring pixels or noise in the classification of high-resolution satellite images is also described as the ‘salt-and-pepper effect’ (Blaschke et al., 2000; Carleer et al., 2005; Hirayama et al., 2018). Hirayama et al. (2018) used a Multiple Classifier System (MCS) approach with six individual machine learning classifiers to reduce the number of isolated pixels by 30 % for all landcover classes and 50 % for the forest class. Alternatively, a pure pixel approach with high-resolution imagery and the delineation of very small high-resolution training areas could have enhanced accuracy in the high-resolution classification. However, wilding pines are very small, individual trees and this would have produced polygons with a small number of total pixels.

Sentinel-2 predicted classification for imagery from 2022

The classifier from the 2015 Sentinel-2 imagery was tested on a 2022 Sentinel-2 scene in the same area. Numerous control operations from boom spraying to manual tree cutting have been conducted over the years which has resulted in the near eradication of wilding pines in this area. In fact, Flock Hill is among the most controlled areas in New Zealand. While in the high-resolution GeoEye-1 imagery no larger group of wildings could be detected by the observers in an RGB and false colour image, the classification of the 2022 Sentinel-2 scene with the classifier from 2015 detected several areas of wilding pines. While the separation of native forests (especially in the northwest of the image but also in the centre) worked well in the 2015 scene, large parts of the native forest in the 2022 imagery were misclassified as wilding pines in the 2022 image.

The ’unhealthy pines’ class achieved reasonably good results, but some of the already grey areas (after several years of control operations) were not detected completely. In summary, applying the 2015 classifier to the 2022 imagery resulted in a modest classification accuracy.

The lack of transferability can partly be attributed to the difference in landcover classes. The most accurately classified class from the 2015 imagery (‘healthy pines’) did not exist in the 2022 scene due to the eradication of wilding pines. Conversely, the worst performing class of ‘unhealthy pines’ was still present in bigger areas, and thus accuracy was comprehensibly low.

Future research should investigate applying the classifier developed from the 2015 imagery to other scenes with existing wilding pines before these have been eradicated. Moreover, future studies could test a multi-temporal approach as described before to improve accuracy in the predicted classification. This could reduce the number of misclassified pixels of wilding pines in the native forest and vice versa. Strategies combining training samples from multiple years and scenes covering a broad range of healthy and unhealthy wilding conifers and other land cover types could also be useful to train a generalizable classifier for Sentinel-2 scenes.

4.2. Assessment of the results of high-resolution WorldView-3 classification

Pixel-based classification results of the high-resolution imagery were substantially lower than for Sentinel-2 with 68 % ground truth accuracy. While the classifiers worked well to separate forest (pines and native forests) from other landcover classes, the distinction between pines and native forests was moderate.

The 2-class classification (Figure 4B) classified the central areas with healthy pines well, but also classified big parts of the native forests and unhealthy trees in the northwest as wilding pines. This poor discrimination between pines and native forests was consistent in all other classification approaches (3- and 4-class classification) which results in a strong overestimation of the pine area. The single acquisition time of the image in summer was probably not ideal for classification, as spectral differences of vegetation are weaker in summer imagery compared to spring or autumn due to changes in phenology (Immitzer et al., 2019; Persson et al., 2018).

The classifier recognised present ‘unhealthy pines’ well but was very sensitive to any dead vegetation like native trees, shrubs or tussock grassland and therefore overestimated the problematic class of ‘unhealthy pines’ in large parts. As with the Sentinel-2 imagery, shadows in steeper areas in combination with rocks and pines were often misclassified as ‘unhealthy pines’. However, the high resolution of the imagery provided very accurate results in terms of delineating vegetation from non-vegetation. Part of the challenge with the ‘unhealthy pines’ class was also that there was no easily defined and clear boundary between dead, nearly dead and starting-to-die trees.

In contrast to the pine area, the area of native forests was underestimated with the high-resolution classifier. Several parts of the native forests were classified as ‘healthy pines’ and some dead and unhealthy trees of the native forest as ‘unhealthy pines’.

Waser et al. (2014) obtained 83% accuracy classifying seven species in a high-resolution WorldView-2 image with an object-based approach. Immitzer et al. (2012) compared an object-based approach with a pixel-based approach for the classification of four and ten tree species in Austria and found that the object-based approach outperformed the pixel-based approach with the best classification accuracy of 95.9 % (four species, 8 bands, object-based) and overall accuracy of 82 % (ten species) compared to a pixel-based accuracy of 88 % (four species) and 73 % (ten species) respectively (Immitzer et al., 2012).

Object-based vs. pixel-based classification

The recognition of ‘unhealthy pines’ was one of the most difficult aspects in this classification both with Sentinel-2 and WorldView-3 imagery for the 2015/2016 imagery (see low accuracies in confusion matrix in Appendix). Especially the pines, which had been killed several years ago, were not detected well as the spectral signature moved away from living vegetation to bare stems with little to no foliage. However, the shape and arrangement of surface spatial structures can also carry landcover information, which is part of object-based classification techniques. The high-resolution imagery pixel-based classification produces small units with a frequent intermixing of different land cover classes which results in less homogenous land cover maps with frequent misclassifications. The spectral differences between conifers and native broadleaved trees could not be distinguished well in the high-resolution imagery with the pixel-based approach. Object-based approaches for the high-resolution dataset might have achieved higher accuracies. Dalponte et al. (2013) classified Norway spruce, Scots pine and Silver birch in Norway with an object-based approach. They concluded that finer spatial resolution significantly increased the classification accuracy of tree species. Similarly, Immitzer et al. (2012) achieved higher accuracy with object-based classification when they assessed the classification of ten tree species with a WorldView-2 image. However, in this study, we worked with very small objects (individual, small trees) and this would have complicated the step of developing candidate objects and producing objects with a small number of total pixels.

The resulting land cover areas in object-based classification are usually more contiguous and homogenous since topology is considered in object-based classification as well. Therefore, more information can be considered in the models of object-based classification and delineation of landcover classes can resemble the perception of an observer more. On the other hand, more homogenous structures also mean a generalisation at the same time and a single or a group of conifers can be harder to detect. Depending on the purpose of the classification, both approaches could be used to either detect single trees or smaller groups or to provide estimates of the extent of bigger infested areas. Hence, the opportunities for landcover classification could potentially be improved when – in addition to pixel-based classification – object-based methods were applied in the high-resolution imagery using both spectral and topological/spatial/geometric features of image objects. For Sentinel-2 data, an object-based approach does not appear to be a promising option because wilding pines tend to be too small for the spatial resolution of 10 m x 10 m and only established and contiguous patches of wildings would be detected.

Classification assessment uncertainty

The spatial analyst toolbox in ArcGIS Pro provided a suitable tool for evaluating users' accuracy with accuracy assessment points. It enhanced the evaluation of the classification results as an additional step to the model validation of the RF model. However, a level of uncertainty remains in the accuracy assessment as it is also a somewhat subjective assessment. In a few cases, it was a challenge for the person performing the accuracy assessment of classified control points to decide what class the points belonged to, especially when points were at the border between two classes or shadows concealed the object of interest. This was especially the case for the 'unhealthy pine' class which was hard to discriminate as the transition from pines, unhealthy pines and tussock or shrubland is in many cases ambiguous. Small (unhealthy) wilding pines versus shrubs or other remnants of native vegetation combined with a mixture of bare soil/slips between wilding pines, or some healthy pines between unhealthy pines made a classification relatively hard in some areas. Steep terrain and resulting shadows can also make classification impossible in some cases.

4.3. Sentinel-2 vs. high-resolution imagery

One of the advantages of using Sentinel-2 data is that the data are available freely, frequently (3-5 days), and at a medium-high resolution (only 10-20 m bands were used in this study). Especially in mountainous areas, where cloud cover can be challenging for analysis of satellite imagery, the short revisit time increases

the likelihood to get cloud-free images. Furthermore, it offers the potential to use images covering all phenological stages throughout the different seasons for more accurate classification.
The more recent imagery (since 2018) is downloadable in a pre-processed format as level 2A data directly as a Bottom-Of-Atmosphere (BOA) corrected reflectance product which makes working with this imagery relatively straightforward (ESA, 2022a).
While individual tree crowns cannot be distinguished with the 10–20 m data, the broad spectral information (with bands in the visible, red-edge, NIR and SWIR wavelengths) have a high potential for tree species separation (Bolyn et al., 2018; Grabska et al., 2019; Grabska et al., 2020; Immitzer et al., 2019; Immitzer et al., 2017; Persson et al., 2018; Puletti et al., 2018), which could also be confirmed in this study. Higher accuracy in tree species classification of Sentinel-2 data compared to high-resolution WorldView-2 data was also discovered by Immitzer et al. (2019), which they attributed to multi-temporal data and the availability of spectral bands in the SWIR spectrum. The consistently high variable importance of red-edge and SWIR bands of Sentinel-2 in this study is in opposition to the variable importance of bands in the WorldView-3 image. WorldView-3 bands and vegetation indices were not as consistent in importance and varied more widely from NDVI, GI, TGI, yellow, green or red-edge bands.

In contrast to Sentinel-2, high-resolution imagery is expensive, and handling and processing take more time, but the higher spatial resolution allows for easier visual interpretation and may be better suited to object-based analysis. Therefore, the potential benefits of high-resolution imagery in terms of accuracy must be weighed against the cost and complexity disadvantages and these considerations become more important the larger the area of interest is. Therefore, the higher costs of the high-resolution imagery should be reflected in significant gains in the accuracy of the landcover classification. However, comparing the classification results of Sentinel-2 and high-resolution imagery in this study it has become clear that Sentinel-2 imagery offers a very suitable, accurate and cost-effective solution to detect wilding pines within a single scene. To work effectively on a large scale across different scenes and acquisition dates future research should focus on the transferability of the classifier to other scenes and points in time.

5. CONCLUSION

In our study, we have investigated the detection of advanced invasive conifer infestations with Sentinel-2 and high-resolution WorldView-3 satellite data using pixel-based classification approaches.
The detection of these exotic conifers with Sentinel-2 imagery worked very well when suitable training areas were delineated in the same imagery. A model accuracy of 99 % and a ground truth accuracy of 96% could be achieved in the classification of ‘healthy pines’. The high accuracy is attributed to the availability of spectral bands in the SWIR and red-edge spectrum and consequently, vegetation indices derived from these bands (RESRI and NBR) were most important. In contrast, pixel-based classification of high-resolution WorldView-3 imagery was less accurate detecting healthy exotic conifers.
The detection of unhealthy, treated conifers was problematic due to the different transition stages of unhealthy-dead trees. They could be detected best during the red-needle stage which occurs shortly after control operations and before needle-fall (grey phase).
The lack of transferability of the classifier from 2015 to a 2022 Sentinel-2 scene was attributed to the complete lack of the most accurately classified class from the 2015 imagery (‘healthy pines’). Conversely, the worst performing class of ‘unhealthy pines’ was present in large areas, and thus accuracy was lower.

Future research should investigate the detection of invasive conifers with a multi-temporal approach to Sentinel-2 imagery. Capturing changes in tree phenology (e.g., leaf flush, leaf colouration, leaf shedding, but particularly changes of colours after control operations) and resulting species-specific changes of the spectral signatures could improve accuracy and reduce misclassification with native trees but also of ‘unhealthy pines’. This could also help to distinguish tree species that otherwise show similar spectral signatures. It might also improve predicted classifications in spatially different areas or later acquisition dates. In addition, applying the classifier developed from the 2015 imagery to other scenes with existing invasive exotic conifers should be explored.

Author Contributions:

This article is based on the master’s thesis written by S.K. at the University of Salzburg. S.K. processed the data, developed the methodology, interpreted the results and wrote the article. B.S. ordered and pre-processed the high-resolution imagery reference data and wrote parts of the corresponding section.

G.P. and T.P. were the supervisors, applied for the funding and contributed to developing the methodology and interpreting the results. Both have reviewed and commented on the article.

Funding:

This research was partly funded by the Vive la résistance research programme (Ministry of Business, Innovation and Employment, New Zealand (MBIE) –Contract C04X2102).

Acknowledgements:

We are grateful to Dr Nicolo Camarretta from Scion for discussions and ideas on the topic; we thank Dr Christian Neuwirth from the University of Salzburg for advice and feedback on a previous version of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

Bannari, A., Morin, D., Bonn, F., & Huete, A. (1995). A review of vegetation indices. *Remote sensing reviews, 13*(1-2), 95-120.

Belgiu, M.; Drăgut, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens., 114, 24–31.

Bivand, R., Keitt, T., Rowlingson, B., Pebesma, E., Sumner, M., Hijmans, R., . . . Bivand, M. R. (2015). Package 'rgdal'. *Bindings for the Geospatial Data Abstraction Library. Available online:* https://cran. r-project. org/web/packages/rgdal/index. html *(accessed on 15 October 2021)*.

Blackburn, G. A. (1998). Spectral indices for estimating photosynthetic pigment concentrations: a test using senescent tree leaves. *International Journal of remote sensing, 19*(4), 657-675.

Blaschke, T., S. Lang, E. Lorup, J. Strobl, and P. Zeil (2000). Object-Oriented Image Processing in an Integrated GIS/Remote Sensing Environment and Perspectives for Environmental Applications. Environmental Information for Planning, Politics and the Public 2: 555–570.

Bolyn, C., Michez, A., Gaucher, P., Lejeune, P., & Bonnet, S. (2018). Forest mapping and species composition using supervised per pixel classification of Sentinel-2 imagery. *Biotechnologie, Agronomie, Société et Environnement, 22*(3), 16.

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

Calle, M. Luz, Urrea, Victor, (2011). Letter to the editor: stability of random Forest importance measures. Brief. Bioinform. 12 (1), 86–89.

Carleer, A. P., O. Debeir, and E. Wolff (2005). Assessment of Very High Spatial Resolution Satellite Image Segmentations. Photogrammetric Engineering and Remote Sensing 71 (11): 1285–1294. doi:10.1117/12.511027.

Criminisi, A.; Konukoglu, E.; Shotton, J. (2012). Decision Forests: A Unified Framework for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning; NOW Publishers: New York, NY, USA.

Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. (2007). Random forests for classification in ecology. Ecology, 88, 2783–2792.

Dalponte, M.; Orka, H.O.; Gobakken, T.; Gianelle, D.; Naesset, E. (2013). Tree Species Classification in Boreal Forests with Hyperspectral Data. IEEE Trans. Geosci. Remote Sens., 51, 2632–2645.

Dash, J. P. (2020). On the detection and monitoring of invasive exotic conifers in New Zealand using remote sensing. *PhD thesis*.

Dash, J. P., Watt, M. S., Paul, T. S., Morgenroth, J., & Hartley, R. (2019a). Taking a closer look at invasive alien plant research: A review of the current state, opportunities, and future directions for UAVs. *Methods in Ecology and Evolution*, *10*(12), 2020-2033.

Dash, J.P., Watt, M.S., Paul, T.S.H., Morgenroth, J. and Pearse, G.D. (2019b). Early Detection of Invasive Exotic Trees Using UAV and Manned Aircraft Multispectral and LiDAR Data. Remote Sensing, 11 (15), 1812.

Dash, J. P., Pearse, G. D., Watt, M. S., & Paul, T. (2017). Combining airborne laser scanning and aerial imagery enhances echo classification for invasive conifer detection. *Remote Sensing, 9*(2), 156.

de Sa, N. C., Carvalho, S., Castro, P., Marchante, E., & Marchante, H. (2017). Using landsat time series to understand how management and disturbances influence the expansion of an invasive tree. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10*(7), 3243-3253.

Dickie, I. A., Sprague, R., Green, J., Peltzer, D. A., Orwin, K., & Sapsford, S. (2022). Applying ecological research to improve long-term outcomes of wilding conifer management. *New Zealand Journal of Ecology*, *46*(2), 1-16.

Einzmann, K., Atzberger, C., Pinnel, N., Glas, C., Böck, S., Seitz, R., & Immitzer, M. (2021). Early detection of spruce vitality loss with hyperspectral data: Results of an experimental study in Bavaria, Germany. *Remote Sensing of Environment*, *266*, 112676.

Elatawneh, A.; Rappl, A.; Rehush, N.; Schneider, T.; Knoke, T. (2013). Forest tree species identification using phenological stages and RapidEye data: A case study in the forest of Freising. In Proceedings of the 5th RESA Workshop, From the Basics to the Service, DLR e.V., Neustrelitz, Germany, 20–21 March 2013; pp. 21–38. 54.

ESA. (2022a). Sentinel-2 MSI User Guide. Retrieved from https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi (accessed on 19/05/2022)

ESA. (2022b). WorldView-3 Instruments. Retrieved from https://earth.esa.int/eogateway/missions/worldview-3 on (accessed on 19/05/2022)

Esri Inc. (2022). *ArcGIS Pro* (Version 2.9.3). Esri Inc. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview

Fassnacht, F.E.; Mangold, D.; Schäfer, J.; Immitzer, M.; Kattenborn, T.; Koch, B.; Latifi, H. (2017). Estimating stand density, biomass and tree species from very high resolution stereo-imagery—Towards an all-in-one sensor for forestry applications? For. Int. J. For. Res., 90, 613–631.

Forkuor, G., Hounkpatin, O. K., Welp, G., & Thiel, M. (2017). High resolution mapping of soil properties using remote sensing variables in south-western Burkina Faso: a comparison of machine learning and multiple linear regression models. *PloS one, 12*(1), e0170478.

Forkuor, G., Dimobe, K., Serme, I., & Tondoh, J. E. (2018). Landsat-8 vs. Sentinel-2: examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. *GIScience & Remote Sensing, 55*(3), 331-354. doi:10.1080/15481603.2017.1370169

Fusaro, L.; Salvatori, E.; Mereu, S.; Manes, F. (2019). Photosynthetic traits as indicators for phenotyping urban and peri-urban forests: A case study in the metropolitan city of Rome. Ecol. Indic., 103, 301–311.

Gitelson, A. A., Buschmann, C., & Lichtenthaler, H. K. (1999). The Chlorophyll Fluorescence Ratio F735/F700 as an Accurate Measure of the Chlorophyll Content in Plants. *Remote Sensing of Environment, 69*(3), 296-302. doi:https://doi.org/10.1016/S0034-4257(99)00023-1

Gitelson, A. A., Kaufman, Y. J., Stark, R., & Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment, 80*(1), 76-87. doi:https://doi.org/10.1016/S0034-4257(01)00289-9

Gitelson, A. A., Merzlyak, M., Zur, Y., Stark, R., & Gritz, U. (2001). Non-destructive and remote sensing techniques for estimation of vegetation status.

Gitelson, A. A., Viña, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., & Leavitt, B. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophysical research letters, 30*(5).

Grabska, E., Frantz, D., & Ostapowicz, K. (2020). Evaluation of machine learning algorithms for forest stand species mapping using Sentinel-2 imagery and environmental data in the Polish Carpathians. *Remote Sensing of Environment, 251*, 112103. doi:https://doi.org/10.1016/j.rse.2020.112103

Grabska, E., Hawryło, P., & Socha, J. (2020). Continuous Detection of Small-Scale Changes in Scots Pine Dominated Stands Using Dense Sentinel-2 Time Series. *Remote Sensing, 12*(8), 1298. Retrieved from https://www.mdpi.com/2072-4292/12/8/1298

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

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

Hastie, T.; Tibshirani, R.; Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed.; Springer: New York, NY, USA; ISBN 978-0-387-84858-7.

Hewitt, A. E. (1998). *New Zealand Soil Classification* (2nd Edition. ed.). Lincoln, New Zealand: Manaaki Whenua Press.

Hijmans, R. J. (2022). Raster: Geographic Data Analysis and Modeling. R package (Version 3.5-15). Retrieved from https://cran.r-project.org/web/packages/raster/

Hill, R.A.; Wilson, A.K.; George, M.; Hinsley, S.A. (2010). Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. Appl. Veg. Sci. 13, 86–99.

Hirayama, H., Sharma, R. C., Tomita, M., & Hara, K. (2019). Evaluating multiple classifier system for the reduction of salt-and-pepper noise in the classification of very-high-resolution satellite images. *International journal of remote sensing, 40*(7), 2542-2557.

Hościło, A.; Lewandowska, A. (2019). Mapping Forest Type and Tree Species on a Regional Scale Using Multi-Temporal Sentinel-2 Data. Remote Sens., 11, 929.

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

Hunt, J. E. R., Doraiswamy, P. C., McMurtrey, J. E., Daughtry, C. S., Perry, E. M., & Akhmedov, B. (2013). A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *International journal of applied earth observation and geoinformation, 21*, 103-112.

Immitzer, M.; Atzberger, C.; Koukal, T. (2012). Tree species classification with Random Forest using very high spatial resolution 8-band WorldView-2 satellite data. Remote Sens., 4, 2661–2693.

Immitzer, M.; Atzberger, C.; Eckert, S. (2017). Assessing the Potential of Sentinel-2 and Pléiades Data for the Detection of Prosopis and Vachellia spp. in Kenya. Remote Sens., 9, 74.

Immitzer, M., Neuwirth, M., Böck, S., Brenner, H., Vuolo, F., & Atzberger, C. (2019). Optimal Input Features for Tree Species Classification in Central Europe Based on Multi-Temporal Sentinel-2 Data. *Remote Sensing, 11*(22), 2599. Retrieved from https://www.mdpi.com/2072-4292/11/22/2599

Kennedy, R.E., Yang, Z., Braaten, J., Copass, C., Antonova, N., Jordan, C. et al. (2015). Attribution of disturbance change agent from Landsat time-series in support of habitat monitoring in the Puget Sound region, USA. Remote Sensing of Environment, 166, 271-285.

Key, C. H., & Benson, N. C. (1999). *Measuring and remote sensing of burn severity.* Paper presented at the Proceedings joint fire science conference and workshop.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., . . . Team, R. C. (2020). Package 'caret'. *The R Journal, 223*, 7.

Landcare Research NZ Ltd (2020a) NZDEM North Island 25 metre, Revision 3. Retrieved 31 October 2021, from https://lris.scinfo.org.nz/layer/48131-nzdem-south-island-25-metre/
Landcare Research NZ Ltd (2020b) Land Environments of New Zealand (LENZ) – Slope, Revision 4. Retrieved 31 October 2021, from https://lris.scinfo.org.nz/layer/48081-lenz-slope/
Landcare Research NZ Ltd (2020c) Land Environments of New Zealand (LENZ) – Mean Annual temperature layer, Revision 4. Retrieved 31 October 2021, from https://lris.scinfo.org.nz/layer/48094-lenz-mean-annual-temperature/
Landcare Research NZ Ltd (2018, Jan. 22) FSL New Zealand Soil Classification, Revision 2. Retrieved 31 October 2021, from https://lris.scinfo.org.nz/layer/48079-fsl-new-zealand-soil-classification/. Soil classification nomenclature from (Hewitt, 1998).
Leckie, D. G., Tinis, S., Nelson, T., Burnett, C., Gougeon, F. A., Cloney, E., & Paradine, D. (2005). Issues in species classification of trees in old growth conifer stands. *Canadian Journal of Remote Sensing*, *31*(2), 175-190.
Leutner, B., Horning, N., Schwalb-Willmann, J., & Hijmans, R. (2019). RStoolbox: Tools for remote sensing data analysis (version 0.2. 4). In.
Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, *2*(3), 18-22.
Lim, J.; Kim, K.-M.; Kim, E.-H.; Jin, R. (2020). Machine Learning for Tree Species Classification Using Sentinel-2 Spectral Information, Crown Texture, and Environmental Variables. *Remote Sens.*, *12*, 2049. https://doi.org/10.3390/rs12122049
Malcolm, J.R.; Brousseau, B.; Jones, T.; Thomas, S.C. (2021). Use of Sentinel-2 Data to Improve Multi-variate Tree Species Composition in a Forest Resource Inventory. Remote Sens., 13, 4297. https://doi.org/10.3390/rs13214297
Mellor, A.; Haywood, A.; Stone, C.; Jones, S. (2013). The performance of random forests in an gperational settingfor large area sclerophyll forest classification. Remote Sens., 5, 2838–2856.
Ministry for the Environment and Statistics New Zealand (2017). Environmental Reporting. Average annual rainfall, 1972–2016, Revision 6. Retrieved 31 October 2021, from https://data.mfe.govt.nz/layer/89421-average-annual-rainfall-19722016/
Ministry for Primary Industries New Zealand. (2014). *The right tree in the right place: New Zealand wilding conifer management strategy 2015-2030. A non-regulatory strategy for the management of wilding conifers in New Zealand.*
Moya, I., & Flexas, J. (2012). Remote sensing of photosynthesis. In J. Flexas, F. Loreto, & H. Medrano (Eds.), *Terrestrial photosynthesis in a changing environment: a molecular, physiological, and ecological approach* (pp. 219-236): Cambridge University Press.
Nagler, P. L., Scott, R. L., Westenburg, C., Cleverly, J. R., Glenn, E. P., & Huete, A. R. (2005). Evapotranspiration on western U.S. rivers estimated using the Enhanced Vegetation Index from MODIS and data from eddy covariance and Bowen ratio flux towers. *Remote Sensing of Environment*, *97*(3), 337-351. doi:https://doi.org/10.1016/j.rse.2005.05.011
Nelson, M. (2017). Evaluating Multitemporal Sentinel-2 Data for Forest Mapping Using Random Forest. Master's Thesis, Stockholm University, Stockholm, Sweden.
New Zealand Forest Owners Association. (2020). New Zealand Plantation Forest Industry Facts & Figures 2019/2020. *NZFO Association (Ed.), 2019/2020*, 35. Retrieved from https://www.nzfoa.org.nz/resources/publications/facts-and-figures
Nuñez, M. A., Chiuffo, M. C., Torres, A., Paul, T., Dimarco, R. D., Raal, P., . . . Van Wilgen, B. W. (2017). Ecology and management of invasive Pinaceae around the world: progress and challenges. *Biological Invasions*, *19*(11), 3099-3120.
Pal, M. (2005). Random forest classifier for remote sensing classification. Int. J. Remote Sens., 26, 217–222.
Pasquarella, V.J.; Holden, C.E.; Woodcock, C.E. (2018). Improved mapping of forest type using spectral-temporal Landsat features. Remote Sens. Environ., 210, 193–207.
Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data, R J., 10, 439–446. In.
Peltzer, D. A. (2018). Ecology and consequences of invasion by non-native (wilding) conifers in New Zealand. *Journal of New Zealand Grasslands*, 39-46.
Peñuelas, J.; Filella, I.; Lloret, P.; Munoz, F.; Vilajeliu, M. (1995). Reflectance assessment of mite effects on apple-trees. Int. J. Remote Sens., 16, 2727–2733.
Persson, M.; Lindberg, E.; Reese, H. (2018). Tree Species Classification with Multi-Temporal Sentinel-2 Data. Remote Sens., 10, 1794.
Pullanagari, R. R., Kereszturi, G., & Yule, I. (2018). Integrating airborne hyperspectral, topographic, and soil data for estimating pasture quality using recursive feature elimination with random forest regression. *Remote Sensing*, *10*(7), 1117.
Puletti, N.; Chianucci, F.; Castaldi, C. (2018). Use of Sentinel-2 for forest classification in Mediterranean environments. Ann. Silvic. Res., 42, 32–38

R Core Team. (2022). R Core Team R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. Retrieved from http://www.R-project.org/
Richardson, D. M., & Rejmánek, M. (2011). Trees and shrubs as invasive alien species–a global review. *Diversity and distributions, 17*(5), 788-809.
Rouse, J., Haas, R., Schell, J., & Deering, D. (1974). Monitoring vegetation systems in the Great Plains with ERTS: Proceedings of the Third Earth Resources Technology Satellite-1 Symposium. *NASA SP-351*, 301-317.
Roy, D. P., Boschetti, L., & Trigg, S. N. (2006). Remote sensing of fire severity: assessing the performance of the normalized burn ratio. *IEEE Geoscience and Remote Sensing Letters, 3*(1), 112-116.
Schindler, J., Dymond, J. R., Wiser, S. K., & Shepherd, J. D. (2021). Method for national mapping spatial extent of southern beech forest using temporal spectral signatures. *International journal of applied earth observation and geoinformation, 102*, 102408. doi:https://doi.org/10.1016/j.jag.2021.102408
Scholten, R. C., Hill, J., Werner, W., Buddenbaum, H., Dash, J. P., Gomez Gallego, M., . . . Watt, M. S. (2019). Hyperspectral VNIR-spectroscopy and imagery as a tool for monitoring herbicide damage in wilding conifers. *Biological Invasions, 21*(11), 3395-3413. doi:10.1007/s10530-019-02055-0
Serrano, L., Ustin, S. L., Roberts, D. A., Gamon, J. A., & Peñuelas, J. (2000). Deriving Water Content of Chaparral Vegetation from AVIRIS Data. *Remote Sensing of Environment, 74*(3), 570-581. doi:https://doi.org/10.1016/S0034-4257(00)00147-4
Smith, R., Adams, J., Stephens, D., & Hick, P. (1995). Forecasting wheat yield in a Mediterranean-type environment from the NOAA satellite. *Australian Journal of Agricultural Research, 46*(1), 113-125.
Sprague, R., Godsoe, W. and Hulme, P.E. 2019 Assessing the utility of aerial imagery to quantify the density, age structure and spatial pattern of alien conifer invasions. Biological Invasions, 21, 2095-2106.
Tong, A., & He, Y. (2017). Estimating and mapping chlorophyll content for a heterogeneous grassland: Comparing prediction power of a suite of vegetation indices across scales between years. *ISPRS journal of photogrammetry and remote sensing, 126*, 146-167.
Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment, 8*(2), 127-150.
Veraverbeke, S., Verstraeten, W. W., Lhermitte, S., & Goossens, R. (2010). Evaluating Landsat Thematic Mapper spectral indices for estimating burn severity of the 2007 Peloponnese wildfires in Greece. *International journal of wildland fire, 19*(5), 558-569.
Visser, V., Langdon, B., Pauchard, A. and Richardson, D. (2014). Unlocking the potential of Google Earth as a tool in invasion science. Biological Invasions, 16 (3), 513-534.
Wardle, P. 1991 Vegetation of New Zealand. Cambridge University Press: Cambridge, 672 p.
Waser, L.T.; Küchler, M.; Jütte, K.; Stampfer, T. (2014). Evaluating the Potential of WorldView-2 Data to Classify Tree Species and Different Levels of Ash Mortality. Remote Sens., 6, 4515–4545.
Wessel, M.; Brandmeier, M.; Tiede, D. (2018). Evaluation of Different Machine Learning Algorithms for Scalable Classification of Tree Types and Tree Species Based on Sentinel-2 Data. Remote Sens., 10, 1419.
Wickham, H. (2016). Data analysis. In *ggplot2* (pp. 189-201): Springer.
Zarco-Tejada, P. J., Miller, J. R., Noland, T. L., Mohammed, G. H., & Sampson, P. H. (2001). Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 39*(7), 1491-1507.
Zhu, Z. (2017). Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. *ISPRS journal of photogrammetry and remote sensing, 130*, 370-384.

APPENDIX

Confusion matrices model accuracy – Sentinel-2

Table 7: Confusion matrix model accuracy assessment for 2-class classification of Sentinel-2 imagery

ClassValue	Other	Healthy pines	Total	U_Accuracy	Kappa
Other	1871	9	1880	100%	0
Healthy pines	0	123	123	100%	0
Total	1871	132		0	0
P_Accuracy	100%	93%	0	99%	0
Kappa	0	0	0	0	96%

Table 8: Confusion matrix model accuracy assessment for 3-class classification of Sentinel-2 imagery

ClassValue	Other	Healthy pines	Unhealthy pines	Total	U_Accuracy	Kappa
Other	1857	10	0	1867	99%	0
Healthy pines	0	122	0	122	100%	0
Unhealthy pines	14	0	72	86	84%	0
Total	1871	132	72		0	0
P_Accuracy	93%	92%	100%	0	99%	0
Kappa	0	0	0	0	0	94%

Table 9: Confusion matrix model accuracy assessment for 4-class classification of Sentinel-2 imagery

ClassValue	Other	Healthy pines	Unhealthy pines	Native forest	Total	U_Accuracy	Kappa
Other	2662	4	1	4	2671	100%	0
Healthy pines	0	205	0	0	205	100%	0
Unhealthy pines	25	0	77	0	102	75%	0
Native forest	129	9	0	1245	1383	90%	0
Total	2816	218	78	1249		0	0
P_Accuracy	95%	94%	99%	100%	0	96%	0
Kappa	0	0	0	0	0	0	92%

Confusion matrices model accuracy – WorldView-3

Table 10: Confusion matrix model accuracy assessment for 2-class classification of WorldView-3 imagery

ClassValue	Other	Healthy pines	Total	U_Accuracy	Kappa
Other	1857	210	2067	90%	0
Healthy pines	383	910	1293	70%	0
Total	2240	1120		0	0
P_Accuracy	83%	81%	0	82%	0
Kappa	0	0	0	0	62%

Table 11: Confusion matrix model accuracy assessment for 3-class classification of WorldView-3 imagery

ClassValue	Other	Healthy pines	Unhealthy pines	Total	U_Accuracy	Kappa
Other	3787	416	277	4480	85%	0
Healthy pines	986	1864	0	2850	65%	0
Unhealthy pines	106	16	1445	1567	92%	0
Total	4879	2296	1722		0	0
P_Accuracy	78%	81%	84%	0	80%	0
Kappa	0	0	0	0	0	67%

Table 12: Confusion matrix model accuracy assessment for 4-class classification of WorldView-3 imagery

ClassValue	Other	Healthy pines	Unhealthy pines	Native forest	Total	U_Accuracy	Kappa
Other	4772	111	352	307	5542	86%	0
Healthy pines	474	3211	5	664	4354	74%	0
Unhealthy pines	95	20	2089	26	2230	94%	0
Native forest	275	402	1	875	1553	56%	0
Total	5616	3744	2447	1872		0	0
P_Accuracy	85%	86%	85%	47%	0	80%	0
Kappa	0	0	0	0	0	0	72%

Ground truthing – Sentinel-2

Table 13: Confusion matrix accuracy assessment for 2-class classification of Sentinel-2 imagery

ClassValue	Other	Healthy pines	Total	U_Accuracy	Kappa
Other	74	4	78	95%	0
Healthy pines	2	61	63	97%	0
Total	76	65	141	0	0
P_Accuracy	97%	94%	0	96%	0
Kappa	0	0	0	0	91%

Table 14: Confusion matrix accuracy assessment for 3-class classification of Sentinel-2 imagery

ClassValue	Other	Healthy pines	Unhealthy pines	Total	U_Accuracy	Kappa
Other	48	2	0	50	96%	0
Healthy pines	2	35	2	39	90%	0
Unhealthy pines	28	0	23	51	45%	0
Total	78	37	25	140	0	0
P_Accuracy	62%	95%	92%	0	76%	0
Kappa	0	0	0	0	0	63%

Table 15: Confusion matrix accuracy assessment for 4-class classification of Sentinel-2 imagery

ClassValue	Other	Healthy pines	Unhealthy pines	Native forest	Total	U_Accuracy	Kappa
Other	35	0	1	1	37	95%	0
Healthy pines	0	28	2	4	34	82%	0
Unhealthy pines	7	1	24	0	32	75%	0
Native forest	2	5	2	25	34	74%	0
Total	44	34	29	30	137	0	0
P_Accuracy	80%	82%	83%	83%	0	82%	0
Kappa	0	0	0	0	0	0	76%

Ground truthing – WorldView-3

Table 16: Confusion matrix accuracy assessment for 2-class classification of WorldView-3 imagery

ClassValue	Other	Healthy pines	Total	U_Accuracy	Kappa
Other	82	2	84	98%	0
Healthy pines	50	26	76	34%	0
Total	132	28	160	0	0
P_Accuracy	62%	93%	0	68%	0
Kappa	0	0	0	0	33%

Table 17: Confusion matrix accuracy assessment for 3-class classification of WorldView-3 imagery

ClassValue	Other	Healthy pines	Unhealthy pines	Total	U_Accuracy	Kappa
Other	46	0	0	46	100%	0
Healthy pines	34	15	0	49	31%	0
Unhealthy pines	16	2	27	45	60%	0
Total	96	17	27	140	0	0
P_Accuracy	48%	88%	100%	0	63%	0
Kappa	0	0	0	0	0	45%

Table 18: Confusion matrix accuracy assessment for 4-class classification of WorldView-3 imagery

ClassValue	Other	Healthy pines	Unhealthy pines	Native forest	Total	U_Accuracy	Kappa
Other	37	1	1	3	42	88%	0
Healthy pines	0	11	2	29	42	26%	0
Unhealthy pines	3	0	20	2	25	80%	0
Native forest	1	1	0	39	41	95%	0
Total	41	13	23	73	150	0	0
P_Accuracy	90%	85%	87%	53%	0	71%	0
Kappa	0	0	0	0	0	0	61%

B) Report

This short technical report describes the methodological steps used in my research in more detail compared to the manuscript. It describes software and functionalities used, pre-processing of Sentinel 1C-data, characterises the machine learning algorithm random forests (RF) among other options and describes the accuracy assessment used for ground truthing.

6.1. Software

This chapter gives an overview of the software I used for this study. Other than the standard Microsoft Office applications I mainly used R and ArcGIS Pro to complete my thesis. For reference purposes I also used SNAP e.g., to compare vegetation indices calculated in R.

6.1.1. R

R is an open-source programming language that is frequently used especially for statistical modelling and analysis. It can handle large datasets and organise them in a structured form. There are many packages for different purposes available and it provides various facilities for carrying out Machine Learning (ML) operations like classification. R is constantly evolving, and large user groups and forums provide good support. Alongside python, this is the most frequently used programming language for geospatial analysis. Sentinel-2 products are downloadable in 100x100 km tiles, thus these large files need to be formatted accordingly.

I have performed all calculations with the satellite data in R version 4.2.0 (R Core Team, 2022). Several packages were used including the 'raster' (Hijmans, 2022), 'RSToolbox (Leutner et al., 2019)', 'rgdal' (Bivand et al., 2015), 'sf' (Pebesma, 2018), 'random forest' (Liaw & Wiener, 2002), and 'caret' (Kuhn et al., 2020) packages and some visualisations using 'ggplot2' (Wickham, 2016).

The functions I used most were from the raster package. These include the following functions:

- compareRaster
- crop
- extent
- hist
- names
- predict
- raster
- resample
- RGB
- shapefile
- stack
- writeRaster

One of the most important functions for the classification was the superClass function from RStoolbox, that was used for the classification of the landcover types. I tested various settings for cross-validation (*k*-fold) and a number of tuning parameter combinations (tuneLength). In the manuscript, I only present the settings which achieved the highest accuracy of the models, (which were a 5-fold cross-validation and a tune length of three). Several other settings have been tested from 1-15. Other packages and functions relevant to the calculations were for example:

- randomForest::random Forest
- randomForest::predict.randomForest
- sf::st_read
- ggplot2:: aes
- rgdal::readOGR
- caret::varImp

6.1.2. ArcGIS

ArcGIS Pro was used for many aspects of my thesis. In general, ArcGIS offers great capabilities for location-based analytics and visualisation and analysis of spatial data. The ArcGIS online version has also been used to create a story map around invasive exotic conifers in New Zealand which is a great tool to promote and visualize a spatially related story or topic.

Most figures and maps of the manuscript were created in ArcGIS. Several of the results calculated in R were checked and visualised in ArcGIS such as the calculated vegetation indices or different bands and bands combinations of the satellite imagery like RGB, false colour etc. Mostly standard geoprocessing and data management tools were used, but I also made use of the Spatial Analyst toolbox e.g., for computation of a confusion matrix and accuracy assessment points. The list of vegetation indices in this study was originally higher than the presented nine indices. Due to some complications and redundancies, which were evaluated in ArcGIS, I dropped a few of these indices.

Furthermore, the site parameters related to growing conditions, including climate (i.e., precipitation, temperature), altitude, slope, or soil type (see manuscript table 1) were retrieved from modelled surface layers from Land Information New Zealand (LINZ) data service.

A very important part of this work was the creation of the training areas for all the different classification approaches and different satellite imagery in ArcGIS. First, the area of interest was defined in which the training areas were located. The training areas were delineated with both Sentinel-2 and high-resolution data but had to be adjusted for the high-resolution due to the distortion in the overlap of the image. The outline of the training areas was always focused to capture the typical spectral signatures (pure pixels) and leave space between the training areas and the borders of the next class. Especially for the Sentinel-2 data, the purity of the pixels was therefore extremely high. Due to the different pixel sizes of Sentinel-2 and WorldView-3, the resulting training areas in a high-resolution Worldview-3 image (1.24 m by 1.24 m resolution) also consist of mixed pixels as also spaces between single wilding pines are included when using basically the same training areas. Using a pure pixel approach for the high-resolution imagery would have complicated this work by producing polygons with a small number of total pixels, as wilding pines are very small, individual trees.

Local knowledge and information about control activities in the area assisted with the delineation. Wilding pine control activities have been conducted over years in the AOI. Most of these measures have been recorded and were made available as a GIS layer for this work to compare what has been treated, when and how. ArcGIS Pro was also used for ground truthing of the classification, which is described in more detail in section 6.4. below.

6.2. Sentinel-2 data pre-processing

The Sentinel-2 satellite data used in this study were downloaded from the Copernicus Open Access Hub **(**https://scihub.copernicus.eu/**).** Since the older Satellite imagery for 2015 was only available as a Level-1C product (Top-Of-Atmosphere product) I used the Sen2Cor package to process it to a Level-2A product for further work. Sen2Cor is a processor for Sentinel-2 products and formatting.

Different versions of Sen2cor exist, but I used version 2.5.5 of Sen2Cor, as this is recommended for older Sentinel-2 Level-1C data products. Using the newer versions Sen2Cor_v2.10 or Sen2Cor_v2.9, which are available at the European Space Agency (ESA) (http://step.esa.int/main/snap-supported-plugins/sen2cor/) website resulted in errors. The pre-processing included atmospheric, terrain and cirrus correction to achieve the result of an orthoimage Bottom-Of-Atmosphere (BOA) corrected reflectance product (ESA, 2022a). It also performs Aerosol Optical Thickness-, Water Vapor-, Scene Classification Maps and Quality Indicators for cloud and snow probabilities (ESA, 2022a). However, cloud and snow cover was unproblematic due to the use of summer scenes with clear sky conditions.

6.3. Classification

Artificial intelligence (AI) helps to solve complex problems with the help of machines, which have become increasingly capable in various fields. Machine Learning (ML) is one part of artificial intelligence, in fact, the dominant technique disclosed in patents (WIPO, 2019), and describes computer algorithms which improve automatically through experience (Russell & Norvig, 2003). A supervised learning approach generally requires a priori knowledge of at least some of the data to help the system determine the statistical criteria for classification. Random Forest (RF) as an ML and supervised learning approach is frequently used in similar research questions around classification (references see manuscript). Other ML approaches that have been used to solve similar problems are the simple and widely used Decision Trees (Gordon et al. 1984), the Support-Vector-Machine (SVM) (Cortes and Vapnik, 1995) or K-Nearest Neighbour (k-NN) (Fix and Hodges, 1951) which was most widely used until the mid-1990s (Russell & Norvig, 2003).

The rationale behind using RF was that it has (A) been successfully used in several other studies, (B) is robust to overfitting and (C) is also robust and effective with unbalanced data (Breiman, 2001; Pal, 2005; Hastie et al., 2009). RF is especially robust to overfitting compared to single decision trees as it basically combines a long list of single decision trees and uses a random subset of variables at each decision node. This is in contrast to a standard decision tree where the nodes of the decision tree are based on the highest decrease in Gini (i.e., the best split). Persson (2018), as well as Belgiu & Drăgut (2016), for example, produced consistently good results compared to other classification algorithms.

The RF model also identifies the relative importance of features during classification which was important comparing several different bands of Sentinel-2 but especially to the spectral bands of WorldView-3 which turned out to be very important. Using this method it was straightforward to find out which of the satellite spectral band values and calculated vegetation indices, were most important as spectral bands and vegetation indices pixels within the training areas were used as predictors in the RF model. The RF algorithm creates an internal unbiased estimate of the generalisation error through the use of bagging (bootstrap aggregation) utilising "out-of-bag" (OOB) samples to produce a diversity of trees (Immitzer et al., 2017). These OBB samples are not part of the training subset (Immitzer et al., 2017). Very detailed explanations about the RF classifier can be found for example in Immitzer et al. (2012) or Immitzer et al. (2017).

However, especially the SVM could have been a suitable alternative to RF and also showed good results in species mapping using Sentinel-2 data (Zagajewski et al., 2021; Grabska et al., 2020).

6.4. Accuracy Assessment

One major goal of this thesis was the comparison of Sentinel-2 and WorldView-3 classifications. Therefore, in addition to the model accuracies and a visual assessment of the classified maps, the ground truthing accuracy is important for quantifying performance independently. Ground truthing means the confirmation or acquisition of knowledge about the study area. This can be attained for instance through fieldwork, analysis of aerial imagery, or personal experience. Information from ground truth is considered to be the most accurate (true) information available about the area of interest (Hexagon, 2012).

The validation of the classification (ground truthing) was performed using accuracy assessment points and confusion matrices. High-resolution imagery as well as local knowledge were used as a reference. Between 130-150, accuracy assessment points outside the training and validation polygons were randomly created in ArcGIS Pro using the create accuracy assessment points geoprocessing tool of the Spatial Analyst Toolbox. The "equalised stratified random" method was selected for the creation of the points. Subsequently, each randomly created accuracy assessment point was assessed for each classification approach (2-4 classes) and satellite data (Sentinel-2 and WorldView-3); "Ground truthing", i.e., determining the correct class for each point, was done using the high-resolution WorldView-3 and GeoEye-1 imagery as a reference and local knowledge of the vegetation. Confusion matrices were computed with the 'compute confusion matrix' tool in the spatial analyst toolbox under 'Segmentation and Classification'. The result of the six different ground truthing assessments can be found in the Appendix of the manuscript (Tables 13-18).

6.5. Author Contributions

The manuscript includes four authors, me as the main author and three other authors. The role(s) of all authors is described as follows:

Sebastian Klinger processed and analysed the data, developed the methodology, interpreted the results and wrote the manuscript article and thesis. Ben Steer ordered and pre-processed the high-resolution imagery reference data and wrote parts of the corresponding section 2.2.2.

Dr Grant Pearse and Dr Thomas Paul were the supervisors of this work at Scion, applied for the funding and contributed to developing the methodology and interpreting the results. Both have reviewed and commented on the manuscript article.

References

Belgiu, M.; Drăgut, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens., 114, 24–31.

Bivand, R., Keitt, T., Rowlingson, B., Pebesma, E., Sumner, M., Hijmans, R., . . . Bivand, M. R. (2015). Package 'rgdal'. *Bindings for the Geospatial Data Abstraction Library. Available online:* https://cran. r-project. org/web/packages/rgdal/index. html *(accessed on 15 October 2021)*.

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

Cortes, Corinna; Vapnik, Vladimir (1995): Support-vector networks. In: Machine Learning 20 (3), S. 273–297. DOI: 10.1023/a:1022627411411.

ESA. (2022a). Sentinel-2 MSI User Guide. Retrieved from https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi (accessed on 19/05/2022)

Esri Inc. (2022). *ArcGIS Pro* (Version 2.9.3). Esri Inc. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview

Fix, Evelyn; Hodges, J. L. (1951): Discriminatory analysis: Nonparametric discrimination: Consistency properties.

Grabska, E., Frantz, D., & Ostapowicz, K. (2020). Evaluation of machine learning algorithms for forest stand species mapping using Sentinel-2 imagery and environmental data in the Polish Carpathians. *Remote Sensing of Environment*, *251*, 112103.

Gordon, A. D.; Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J. (1984): Classification and Regression Trees. In: Biometrics 40 (3), S. 874-889. DOI: 10.2307/2530946.

Hastie, T.; Tibshirani, R.; Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed.; Springer: New York, NY, USA; ISBN 978-0-387-84858-7.

Hexagon (2012). Erdas Imagine Producer Field Guide. Retrieved from https://hexagongeospatial.fluidtopics.net/r/uOKHREQkd_XR9iPo9Y_Ijw/khBQoaH4ne8R6PVMWZ7Ubw accessed on 21.01.2022

Hijmans, R. J. (2022). Raster: Geographic Data Analysis and Modeling. R package (Version 3.5-15). Retrieved from https://cran.r-project.org/web/packages/raster/

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

Immitzer, M.; Atzberger, C.; Koukal, T. (2012). Tree species classification with Random Forest using very high spatial resolution 8-band WorldView-2 satellite data. Remote Sens., 4, 2661–2693.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., . . . Team, R. C. (2020). Package 'caret'. *The R Journal*, *223*, 7.

Leutner, B., Horning, N., Schwalb-Willmann, J., & Hijmans, R. (2019). RStoolbox: Tools for remote sensing data analysis (version 0.2. 4). In.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, *2*(3), 18-22.

Pal, M. (2005). Random forest classifier for remote sensing classification. Int. J. Remote Sens., 26, 217–222.

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data, R J., 10, 439–446. In.

Persson, M.; Lindberg, E.; Reese, H. (2018). Tree Species Classification with Multi-Temporal Sentinel-2 Data. Remote Sens., 10, 1794.

R Core Team. (2022). R Core Team R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. Retrieved from http://www.R-project.org/

Russell, Stuart J.; Norvig, Peter (2003). Artificial Intelligence: A Modern Approach (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall, ISBN 0-13-790395-2.

WIPO (World Intellectual Property Organisation) (2019). WIPO Technology Trends 2019 – Artificial Intelligence. Retrieved from: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf accessed on 27/09/2022.

Zagajewski, B.; Kluczek, M.; Raczko, E.; Njegovec, A.; Dabija, A.; Kycko, M. (2021). Comparison of Random Forest, Support Vector Machines, and Neural Networks for Post-Disaster Forest Species Mapping of the Krkonoše/Karkonosze Transboundary Biosphere Reserve. Remote Sens., 13, 2581. https://doi.org/10.3390/rs13132581

Wickham, H. (2016). Data analysis. In *ggplot2* (pp. 189-201): Springer.