Mine Closure 2025 – S Knutsson, AB Fourie & M Tibbett (eds)

doi:10.36487/ACG_repo/2515_67

Monitoring success: remote sensing and artificial intelligence for tracking success of individual seedlings at a revegetation trial in northern Canada

Jeff Anderson [a,*], Ben Pearse [a]

[a] *Integral Ecology Group, Canada*

Abstract

*Progressive reclamation and reclamation trials are vital to mine closure success, providing opportunities to refine and validate reclamation prescriptions. Comprehensive monitoring is essential for evaluating the relative success of treatments over time. The research objective was to identify newly planted tree seedlings, classify them by species, and develop a robust indicator of seedling health using high-spatial-resolution multispectral imagery, object-based image analysis, and machine learning. In July 2023, a 2 hectare revegetation trial established on a mined waste rock dump was planted with aspen (*Populus tremuloides*), lodgepole pine (*Pinus contorta*), and white spruce (*Picea glauca*). Multispectral imagery was acquired via remotely piloted aircraft system (RPAS) in August 2023, and 30 permanent sampling plots (PSPs) were established across the trial, where each seedling's position and species were recorded (n = 1,545). The multispectral imagery was filtered using a vegetation index to isolate individual seedlings across the trial, and a Random Forest model was used to classify seedlings by species. Subsequent RPAS imagery and vigour rankings for seedlings within PSPs were collected in August 2024 and analysed using a similar approach to classify survival (dead or alive) and vigour (0–5) or generate seedling size percentiles as indicators of health. Species and survival were classified with an average F1-score of 0.94 and 0.98, respectively. Vigour classifications for aspen, lodgepole pine, and white spruce yielded average F1-scores of 0.82, 0.86, and 0.83, respectively. Increases in size percentiles coincided with increases in field-assessed vigour, demonstrating a clear relationship between the two. This study resulted in the precise geotagging of 4,942 aspen, 2,524 white spruce, and 2,819 lodgepole pine across the reclamation trial, demonstrating the ability to monitor the progression of individual seedling performance immediately after planting. This approach improves the efficiency of monitoring planting programs, enabling deeper insights into the long-term effects of different treatments on seedlings.*

Keywords: *revegetation trial, drone, UAV, RPAS, remote sensing, machine learning*

1 Introduction

Revegetation efforts in northern environments face a unique set of challenges, most notably short growing seasons and harsh winters, which naturally slow down the process of ecological recovery (Deshaies et al. 2009). Mine sites often present additional difficulties such as steep slopes, soils that are naturally unfertile with low capacities for water and nutrient retention, fluctuations in both moisture and temperature, acidic soil conditions, and the potential for significant heavy metal contamination, all of which can have different impacts on the establishment and health of different species of vegetation (Asmara et al. 2022; Tordoff et al. 2000). This myriad of factors can make it difficult to separate the effects of environmental conditions from the effects of different management approaches to reclamation. Robust monitoring of revegetation efforts is required to disentangle these factors to inform future reclamation and to understand the likely outcomes and timelines that can be expected with respect to the restoration of mine sites.

* Corresponding author. Email address: jkanderson@outlook.com

Traditional vegetation monitoring methods typically involve establishing permanent sampling plots (PSPs) to collect and track vegetation data over time. These methods can struggle to capture spatial nuances of trial areas and rely on the assumption that reclamation treatments are somewhat homogenous. These methods also require frequent visits to sensitive ecosystems and personnel hours close to potentially active mining areas.

This study presents a fusion of remote sensing, artificial intelligence and traditional monitoring methods to maximise learnings from a trial established as part of progressive reclamation of a mine site, while reducing the ground-based sampling efforts.

1.1 Remotely piloted aircraft systems for mine reclamation monitoring

The advent of remote sensing technologies has brought significant advancements in ecological monitoring and vegetation assessment, offering solutions to some limitations of traditional methods. Remote sensing provides a practical and economical means to study changes in vegetation cover and characteristics, especially over large and often inaccessible areas, thereby overcoming many of the accuracy and coverage limitations associated with ground-based approaches (Manfreda et al. 2018). While satellites are commonly used to assess the characteristics of vegetated ecosystems at the regional, continental and even global scale (Xie et al. 2008), remotely piloted aircraft systems (RPAS), commonly known as drones or unmanned aerial vehicles, offer a solution for collecting very high-resolution data over a site or landscape level (Manfreda et al. 2018).

Equipped with a variety of sensors including true-colour optical, multispectral optical, and LiDAR, RPAS can capture high-spatial-resolution data which are well-suited for detailed vegetation assessment (Anderson et al. 2021). When paired with centimetre-accurate georeferencing, these datasets can be accurately and easily correlated with ground-based measurements collected at comparable accuracy (Aasen et al. 2018). Potential applications of RPAS-based vegetation assessment include monitoring plant populations, assessing biodiversity and species richness (Bagaram et al. 2018; Fernández et al. 2020; Saarinen et al. 2018), detecting the invasion of non-native plant species (Álvarez-Taboada et al. 2017), evaluating the effectiveness of restoration efforts (Anderson et al. 2021; Meng et al. 2024; Reinprecht and Kieffer 2025), tracking plant phenology (Klosterman et al. 2018), identifying pest infestations in forests (Vanegas et al. 2018), observing changes in land cover, and retrieving valuable structural information about vegetation for forest assessment and management (Hologa et al. 2021; Manfreda et al. 2018; Xie et al. 2008).

Operationally, the inherent flexibility and rapid deployment capabilities of RPAS make them well-suited for monitoring reclamation sites, where remote data collection can reduce vehicle and foot-traffic in ecologically sensitive areas, and the safety risks of personnel time in active mining areas. Unlike satellite imagery, which can be affected by cloud cover and typically has fixed revisit schedules (Iizuka et al. 2018), RPAS can be deployed on demand, providing timely data for monitoring specific events or for assessing the immediate impacts of environmental factors.

1.2 Artificial intelligence for remote sensing analysis

Integration of artificial intelligence (AI) and machine learning, a subset of AI, into the analysis of remote sensing data can make the processing and analysis steps much more efficient and further increase the value-proposition of remote sensing integration with reclamation monitoring programs. Machine learning algorithms offer a powerful approach for analysing complex remote sensing datasets across various environmental fields, including the critical area of vegetation monitoring (Janga et al. 2023; Maxwell & Fang 2018). The application of AI to remote sensing data enables the automation of intricate analytical tasks, leading to increased efficiency, enhanced objectivity in the analysis, and the ability to process vast datasets that would be practically impossible to analyse through manual methods alone. By learning from data collected during ground-based surveys, machine learning models can be developed to expand findings from the plot level to the extent of the data collected during the RPAS survey.

The combination of RPAS-collected imagery, object-based image analysis (OBIA), and machine learning classifications has the potential to extrapolate findings from ground-based surveys at sample plots to the entire study area. OBIA is an image analysis approach that moves beyond traditional pixel-based classification by grouping spectrally similar and spatially contiguous pixels into meaningful image objects (Blaschke et al. 2014). This object-based approach allows for the incorporation of not only spectral information but also spatial and contextual information, such as the shape, texture, and spatial relationships between different image objects, which can be crucial for accurate vegetation classification, especially when using high-resolution imagery where individual plants or small groups of plants can be delineated as distinct entities.

The Random Forest algorithm (Breiman 2001) is a widely used machine learning method with proven efficiency and robustness in classifying large and complex datasets, including multispectral imagery acquired from various remote sensing platforms (Belgiu & Drăguţ 2016; Homer et al. 2004). Random Forest operates by constructing an ensemble of decision trees during the training phase and then aggregating their predictions through majority voting to determine the final classification of new data points. This approach provides a method for classification that is generally less sensitive to noise in the data and less prone to overfitting (where the model is very efficient at classifying the training data but struggles to classify new, unseen data), making it particularly well-suited for ecological applications (Belgiu & Drăguţ 2016).

The combination of OBIA and Random Forest leverages the strengths of both approaches. OBIA provides the framework for identifying individual plants or seedlings as discrete objects within the imagery, while Random Forest offers a method for classifying these objects into different species or for assessing their health based on their spectral and spatial attributes. Numerous studies have demonstrated the effectiveness of OBIA and Random Forest in a variety of vegetation mapping and monitoring applications, including the accurate identification and classification of different tree species (Fassnacht et al. 2016; Hologa et al. 2021; Immitzer et al. 2012) and the detailed assessment of vegetation health (Guerra-Hernández et al. 2021; Näsi et al. 2015) from high-resolution remote sensing data; however, few are performed on tree seedlings less than 50 cm in height.

1.3 Monitoring success in planting trials

Assessing the growth and survival of seedlings planted in reclamation trials and progressive reclamation often involves tagging seedlings directly, which can damage the seedlings, tagging a wire post planted next to the seedlings, which is expensive and time-consuming to install, or planting in particular patterns which is difficult to scale operationally. Each of these methods usually result in lost tags or seedlings, reducing the number of seedlings which can be included in the analysis. Combining high-accuracy geolocation, high-spatial-resolution RPAS imagery, OBIA and machine learning analysis, as described above, provides an opportunity to identify and track all seedlings planted within a planting trial efficiently and effectively. By assessing individual seedlings across an entire planting area, researchers may be able to disentangle the effects of various treatments on seedling survival and health, while compensating for the spatial effects of different environmental conditions across the study area. This isn't possible when monitoring vegetation at coarser spatial scales, such as at the plot level.

This study addresses the need for precise and efficient monitoring of individual seedling performance in a revegetation trial conducted on a mined waste rock dump in northern Canada. By employing the methods introduced above, this research aims to identify individual seedlings, classify them by species, including aspen (*Populus tremuloides*), lodgepole pine (*Pinus contorta*), and white spruce (*Picea glauca*), and develop a robust indicator of seedling health. Together, these efforts establish the foundation for a remote sensing-based monitoring program capable of tracking seedling health responses to various reclamation treatments. This study demonstrates the potential of remote sensing and AI for improving the efficiency and insights gained from monitoring mine site revegetation programs.

2 Methods

2.1 Site description

A 2 hectare revegetation trial was established on a mined waste rock dump, comprising two distinct areas: a plateau (~1 hectare) divided into 19 panels (~0.05 hectares each) and a sloped area (~1 hectare) divided into five panels (~0.18 hectares each) (Figure 1). The final trial as-built design contains the following key variables:

- Materials: two types sourced from separate borrow areas. Borrow Area 8 (BA8) material is a fine-grained glacial till. Borrow Area 17E (BA17E) material, although understood to be of similar glacial origin to BA8, is more coarse-grained, with a weathered/oxidised appearance (i.e. BA17E till is a light-brown colour, while BA8 till is grey).
- Depth treatments: BA8 material was placed at three depths: 0.3, 0.6, and 1.0 m. The BA17 was placed at a depth of 1.0 m for all panels.
- Material placement and site preparation methods: a variety of material placement and site preparation methods were implemented, resulting in panels with varying bulk densities. For example, high bulk densities were achieved through equipment trafficking, where materials were placed and compacted by a dozer. Low bulk densities were achieved by placing material with an excavator, mounding techniques, and loosening the surface of panels (following methods popularised by Dave Polster [Polster 2016]).

In June 2023, all panels were planted with three tree species: white spruce (*Picea glauca*), lodgepole pine (*Pinus contorta*), and aspen (*Populus tremuloides*). Additionally, the slope panels were seeded at a rate of 30 kilograms per hectare with grass seed to prevent erosion.

Figure 1 Study area, showing panel layout with permanent sampling pots

2.2 Ground-based surveys

In June 2023, immediately after planting, 30 PSP (1 plot per panel on the plateau and two plots per panel on the slopes) were established across the trial. All tree seedlings within each PSP were geolocated using a

Stonex S900A GNSS receiver (~0.5 cm accuracy) and attributed with a species ID. Additionally, clumps of grasses which had emerged in August 2023 were geolocated for use as training data to separate seedlings from grasses in classification models.

In 2024, these seedlings were revisited, and each was assigned a vigour score. Vigour of each seedling was ranked from 0 to 5 based on a modified needle-colour index (Berquist et al. 2003), where 0 indicates needles are 100% brown and 5 indicates needles are 100% green. This methodology was adapted for broadleaf by benchmarking relative foliage amount to the best-performing seedling of that species. The criteria for vigour rankings are presented in Table 1. For the purposes of analysis, ranks 0 and 1 are dead (0) and dying (1), and are pooled as dead for the purpose of classifying seedlings as alive or dead.

Table 1 Criteria for qualitative assessment of seedling vigour

Vigour ranking	Needle-colour index for conifer seedlings	Adapted foliage fullness for broadleaf seedlings
0	100% brown or missing	No leaves or missing
1	~90% of needles brown	~5% or less maximum foliage
2	~50% of needles brown, 50% green	~5–35% of maximum foliage
3	~75% of needles green, 25% brown or yellow	~35–65% of maximum foliage
4	~90% needles green, 10% yellow	~65–95% of maximum foliage
5	100% of needles green	~95% or higher maximum foliage

2.3 Remotely piloted aircraft system data collection and post-processing

Optical imagery was collected using a DJI Matrice 300 equipped with a Micasense Altum and DJI Zenmuse P1 during August 2023 and 2024 (Flight settings shown in Tables 2 and 3). The Micasense Altum collects multispectral imagery in the blue, green, red, red edge, near-infrared and shortwave infrared spectral bands. It is fitted with a downwelling light sensor fixed to the top of the M300 for measuring incident irradiance and is used for correcting within-flight changes in illumination conditions (Daniels et al. 2023). The DJI Zenmuse P1 collects true-colour (red, green, and blue) images.

Imagery collected in August 2023 was used to support species classification. The timing of image collection at the end of the growing season is important, as aspen trees had not developed leaves in June 2023. Seedling heights during this data collection period ranged from approximately 5–30 cm. Imagery collected in August 2024 was used to evaluate seedling health following two summers of growth and one winter season. Seedling heights during this data collection period ranged from approximately 5–50 cm.

Table 2 Flight settings for image collection using the Micasense Altum

Flight parameter	Setting value
Side overlap	80%
Front overlap	85%
Gimbal angle	nadir
Ground sample distance	2.0 cm/pixel
Altitude	44 m above ground level
Speed	4.5 m/s

Table 3 Flight settings for image collection using the DJI Zenmuse P1

Flight parameter	Setting value
Side overlap	80%
Front overlap	80%
Gimbal angle	Smart oblique
Ground sample distance	1.74 cm/pixel
Altitude	100 m above ground level
Speed	5 m/s

For each flight, ground control points (GCPs) were distributed throughout the study area and a GNSS receiver (~0.5 cm accuracy) was used to record their locations. Ground sample distance (GSD) and altitude will vary slightly from those presented in Tables 2 and 3. In 2024, terrain follow was used, keeping the RPAS at a consistent height above the plateau and sloped areas. Terrain follow was not used in 2023, meaning that areas near the bottom of the slope will have a slightly larger GSD.

All imagery was processed in Agisoft Metashape following the methods of Śledź & Ewertowski (2022). GCPs collected in the field were used for georeferencing. Images from the Micasense Altum were radiometrically normalised using the empirical line method, which uses a reflectance panel to convert raw pixel values into reflectance (Aasen et al. 2018). The final outputs from this step were: a multispectral and true-colour orthomosaic from August 2023, and a multispectral and true-colour orthomosaic from August 2024.

2.4 Object-based image analysis and supervised Random Forest classifications

2.4.1 Isolation and attribution of vegetation clusters

In this study, we employed a simplified form of OBIA, using spectral thresholding of a vegetation index to delineate discrete objects representing individual seedlings. While traditional OBIA workflows often rely on more complex segmentation algorithms, our threshold-based approach achieved a similar outcome: the generation of meaningful vegetation objects suitable for object-level analysis.

For each of the vegetation assessment periods (August 2023 and August 2024), a modified soil-adjusted vegetation index (MSAVI) raster was generated using the multispectral orthomosaics. The MSAVI vegetation index was used for isolation of vegetation clusters because it minimises contributions of background bare soil to the index value (Pan et al. 2023), this is particularly important for this study, given the different colouration of the two borrow materials being assessed.

An MSAVI filter value was developed and used to identify and isolate vegetation clusters. The filter value was determined using a process of manually checking MSAVI values at areas of known vegetation, known soil, and known non-vegetation areas, which may have high MSAVI values (such as woody debris, flagging tape and shadow areas). For each of the assessment periods, a filter value of 0.1 was appropriate to reliably capture vegetation, without capturing non-vegetation areas. All vegetation clusters were then polygonised into image objects.

Isolation of vegetation polygons was successful for all panels using the August 2023 imagery (n= 14,430). Using the 2024 imagery, vegetation polygons were successfully isolated for plateau panels (n = 8,616), but not for the slope panels. Grass seeded on the slope panels at the time of tree planting was sufficiently established to prohibit isolation of seedlings.

Each polygon, for each year, was then attributed with an array of spectral and geometric characteristics which would be used as factors in the classification model. This was done using the Rasterio (Gillies 2019) and Geopandas (Jordahl et al. 2020) packages in Python. The characteristics included 74 spectral characteristics (mean, 75th, 90th, 95th, 99th, 100th percentiles of blue, green, red, red edge, near-infrared, NDVI, MSAVI, NDRE, RVI, GRNVI and GNDVI), count of pixels above MSAVI thresholds (0.1, 0.2, 0.4, 0.6, and 0.8) and four geometric characteristics (area, perimeter, aspect ratio, and compactness).

The dataset was finalised by performing a spatial join between the seedlings within PSPs that were geolocated and attributed with a species ID in 2023 and the vegetation polygons. This served to assign a seedling species ID to a subset of polygons (n = 2,346) to be used as species training data in the subsequent classifications. Vegetation polygons within PSPs which were not identified as planted seedlings were used as training data for non-seedling vegetation. The vigour score as assessed from the ground-based surveys of seedlings in PSPs, was joined with the 2024 seedling polygons.

2.4.2 Random Forest model development and assessment

The Random Forest classifier module available in the Scikit-learn Python package (Pedregosa et al. 2018) was used to train and apply each of the models. All models were trained with a train-test division of 80/20, meaning that for each model, 80% of the data was used for training and 20% was held out for independent testing. This enables final model performance to be evaluated by making predictions on unseen data and comparing those predictions to the known class labels. Feature selection was employed using recursive feature elimination with cross-validation (RFECV) in the Scikit-learn package (Pedregosa et al. 2018). Imbalanced classes within datasets were artificially balanced using the synthetic minority oversampling technique (SMOTE) in the imbalanced-learn package (Lemaître et al. 2017).

Accuracy reports of models were generated using the 'classification_report' function in Scikit-learn. The F1-score was used to describe performance of each class, where:

$$F1 = \frac{2*True\ Positives}{2*True\ Positives+False\ Positives+False\ Negatives} \tag{1}$$

Furthermore, the average-F1 score was used to evaluate how well the model performed across the entire dataset.

A classification score (obtained using the predict_proba function in Scikit-learn) was used to further refine classifications. For each seedling, this score represents the proportion of decision trees in the Random Forest that voted for the predicted class (aspen, lodgepole pine, or white spruce), calculated by dividing the number of trees that voted for the majority class by the total number of trees in the model. Where possible, classifications were manually checked using the 2024 optical imagery, where any seedling with a classification score below 0.7 was checked and edited where necessary. This was only possible for some classifications (i.e. survival, species) and where a confident assessment could be made from the imagery.

2.5 Seedling species classification

Seedling species classifications were attempted using the imagery collected in August 2023 and 2024 at a time when aspen had leafed out, intraspecies spectral variation was close to its minimum, and interspecies spectral variation was near its maximum.

Using the 2023 dataset, a model was trained to classify the vegetation polygons into grass and seedlings. It was then applied to classify all vegetation polygons into these two categories. This was not repeated using the 2024 dataset, as the seedlings could not be isolated from grasses on the slope panels.

The datasets were then used to train models classify the seedling polygons into species (only the seedling polygons on the plateau panels were used from the 2024 dataset). The model was then used to assign a species classification and associated classification score to all seedling polygons.

The final dataset produced in this step consists of seedling polygons, each assigned a species ID, covering the entire study area. The classifications from the 2023 dataset provided significantly better results (Section 3.1), which applied to the entire study area, so the 2023 dataset was used to assign a species ID to the seedling polygons in the 2024 dataset for evaluating seedling health in the subsequent steps.

2.6 Identifying a suitable seedling health indicator

While the seedling species classifications performed in the previous step enabled individual seedlings to be monitored by species, a suitable indicator of seedling health still needed to be selected. This section evaluated three potential indicators:

1. Random Forest classification of seedling survival.
2. Random Forest classification of seedling vigour based on field-estimated ratings.
3. Seedling size percentiles derived based on the area of individual seedlings.

Only the 2024 dataset was used for monitoring seedling health, as differentiation in health between seedlings was not expected after only two months from the time of planting in the 2023 dataset. Therefore, the health of individual seedlings could only be assessed for those planted in the plateau panels.

2.6.1 Classification of seedling survival and vigour using object-based image analysis and a Random Forest model

The vigour score of seedlings within the PSPs were used for training a model to classify seedling survival, where vigour scores of 0 and 1 were pooled as ‘dead’ and scores of 2, 3, 4 and 5 were pooled as ‘alive’. Using the Random Forest algorithm and method described above, the survival status of each 2024 seedling on the plateau was classified.

Similarly, the vigour score of seedlings within the PSPs were used for training a model to classify seedling vigour, where vigour scores of 0 and 1 were pooled as ‘dead’, scores of 2 and 3 were pooled as ‘moderate’, and scores of 4 and 5 were pooled as ‘high’. Using the Random Forest algorithm and method as described above, the vigour pool of each 2024 seedling on the plateau was classified. The model was then used to assign a vigour classification and associated classification score to all seedling polygons on the plateau.

2.6.2 Estimation of seedling health using seedling size percentiles

Seedling size was calculated for each seedling by using a count of pixels above the MSAVI vegetation filter within each seedling polygon and multiplying the count by the imagery’s spatial resolution (in cm^2 per pixel). Each seedling was then scored a percentile (scaled from 0–100) of its size relative to the distribution of sizes for seedlings of that species. The validity of this approach was confirmed by comparing the seedling percentiles to the vigour assessments collected in the field.

3 Results

3.1 Seedling species classification

The Random Forest model used to classify the vegetation polygons into grass and seedlings achieved an average F1-score of 0.94. Following classification, a robust quality check was performed after this step, where all classifications with a confidence score of 0.7 or lower were visually confirmed (n = 622), and 55 misclassifications were identified and corrected.

The Random Forest model used to classify tree species achieved strong performance with an average F1-score of 0.88 (Table 4). The model performance was highest for aspen, with an F1-score of 0.96, while lodgepole pine and white spruce had F1-scores of 0.85 and 0.83, respectively.

Lower F1-scores for lodgepole pine and white spruce are because these two species are sometimes misclassified as each other. This is likely due to these two species sharing similar spectral and structural characteristics. Notably, area was an important feature in the classification model, and lodgepole pine and white spruce seedlings are more similar in size compared to aspen.

Table 4 F1-scores for individual species in the classification model

Species	F1-score
Aspen	0.96
Lodgepole pine	0.85
White spruce	0.83
Average	0.88

While the species classification model demonstrated strong overall performance, manual edits on seedlings with a classification score below 0.7 using high-resolution, true-colour imagery were effective for further increasing classification accuracies.

The Random Forest model used to classify species using the August 2024 dataset achieved an average F1-score of 0.77, a consequence of the wide range in seedling vigour. These results can be improved (to 0.81) by filtering the seedlings classified as dead (Section 3.2.1).

3.2 Identifying a suitable seedling health indicator

3.2.1 Random Forest classification of survival and vigour

The Random Forest model trained for classification of the survival status of seedlings showed excellent performance, achieving an average F1-score of 0.96 (Table 5**Error! Reference source not found.**). Model performance was highest for alive seedlings, with an F1-score of 0.99, while dead seedlings were classified with an F1-score of 0.92.

The lower F1-score for dead seedlings was the result of four seedlings being classified as dead when they were alive. A limitation of this classification was the limited number of dead seedlings used in the training data (n = 22), which can challenge model generalisation for minority classes. The SMOTE was applied to address this imbalance, however, the small sample size may still have limited the model's ability to fully capture the variability within the dead seedling class. Despite this, the high overall performance indicates that the model is reliable for distinguishing seedling survival across the dataset.

Table 5 F1-scores for dead and alive seedlings in the classification model

Species	F1-score
Alive	0.99
Dead	0.92
Average	0.96

The Random Forest model used to classify seedling vigour achieved similar levels of performance across species and vigour classes. Aspen, lodgepole pine, and white spruce achieved average F1-scores of 0.82, 0.86, and 0.83, respectively (Table 6). Across all species, the model accurately identified low vigour seedlings. However, it showed reduced performance when classifying moderate and high vigour classes. F1-scores for high vigour classes were 0.67, 0.76, and 0.56, for aspen, lodgepole pine, and white spruce, respectively, indicating greater confusion between these upper vigour categories.

Across all species, the main source of classification error was the result of confusion between the moderate and high vigour classes. Seedlings in the low class, characterised by needles or leaves that were almost entirely brown, exhibited a distinct spectral signature, making them relatively easy to separate from the other classes. In contrast, moderate and high vigour seedlings were less spectrally distinct, as both had green foliage and were potentially reflecting near-infrared (NIR) similarly. This spectral overlap, and lack of a distinct separation between the two classes likely contributed to the misclassifications.

While the results of the vigour classifications showed moderate overall performance, the variability in performance across vigour classes was sufficient enough to warrant exploration of an alternative approach.

Table 6 F1-scores for individual vigour classes for each species in the classification model

Species	Class	F1-score
Aspen	Low	1
	Moderate	0.81
	High	0.67
	Average	0.82
Lodgepole pine	Low	0.92
	Moderate	0.89
	High	0.76
	Average	0.86
White spruce	Low	1
	Moderate	0.93
	High	0.56
	Average	0.83

3.2.2 Estimation of seedling health using seedling size percentiles

Figure 2 shows a clear relationship between field-assessed vigour classes and seedling size percentiles. Seedlings assigned lower vigour classes (0 and 1) consistently fall within the lowest size percentiles and increases in field-assessed vigour coincide with a steady increase in size percentiles, demonstrating that larger seedling size is generally associated with higher vigour. This alignment between field observations and remotely sensed size metrics supports the validity of using MSAVI-derived size percentiles as a proxy for seedling health.

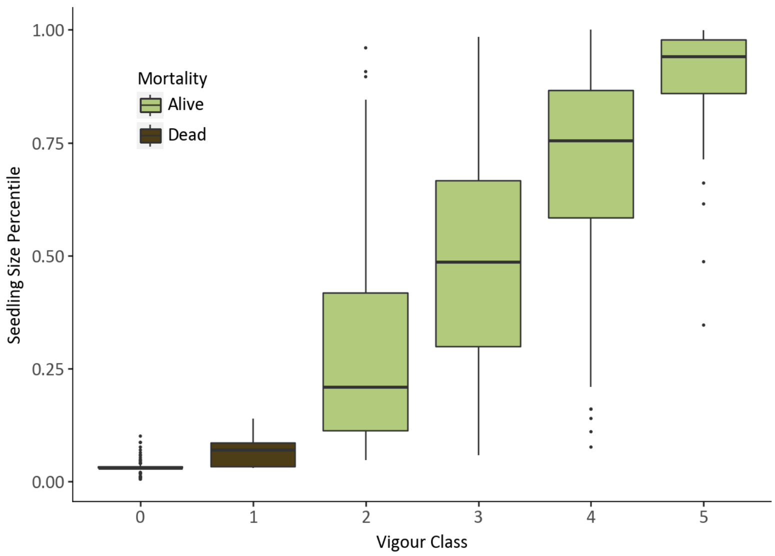

Figure 2 Boxplots showing the relationship between seedling vigour class and seedling size percentile. Vigour classes were assigned through field-based assessments, while seedling size percentiles were derived using modified soil-adjusted vegetation index thresholding and pixel-based area calculations within seedling polygons

This approach enables the comparison of seedling health within each species, as each seedling's size percentile reflects its relative size among conspecifics. However, because size percentiles were calculated separately for each species, they cannot be used to compare seedling health between species. For example, an aspen seedling in the 90th percentile is large relative to other aspen seedlings, but this does not imply it is larger or healthier than a pine seedling in the 70th percentile of its group.

Combined with the Random Forest classification of seedling species, seedling size percentiles can be used to map the spatial distribution of both species and seedling health across the study area, enabling consistent monitoring over time (Figure 3).

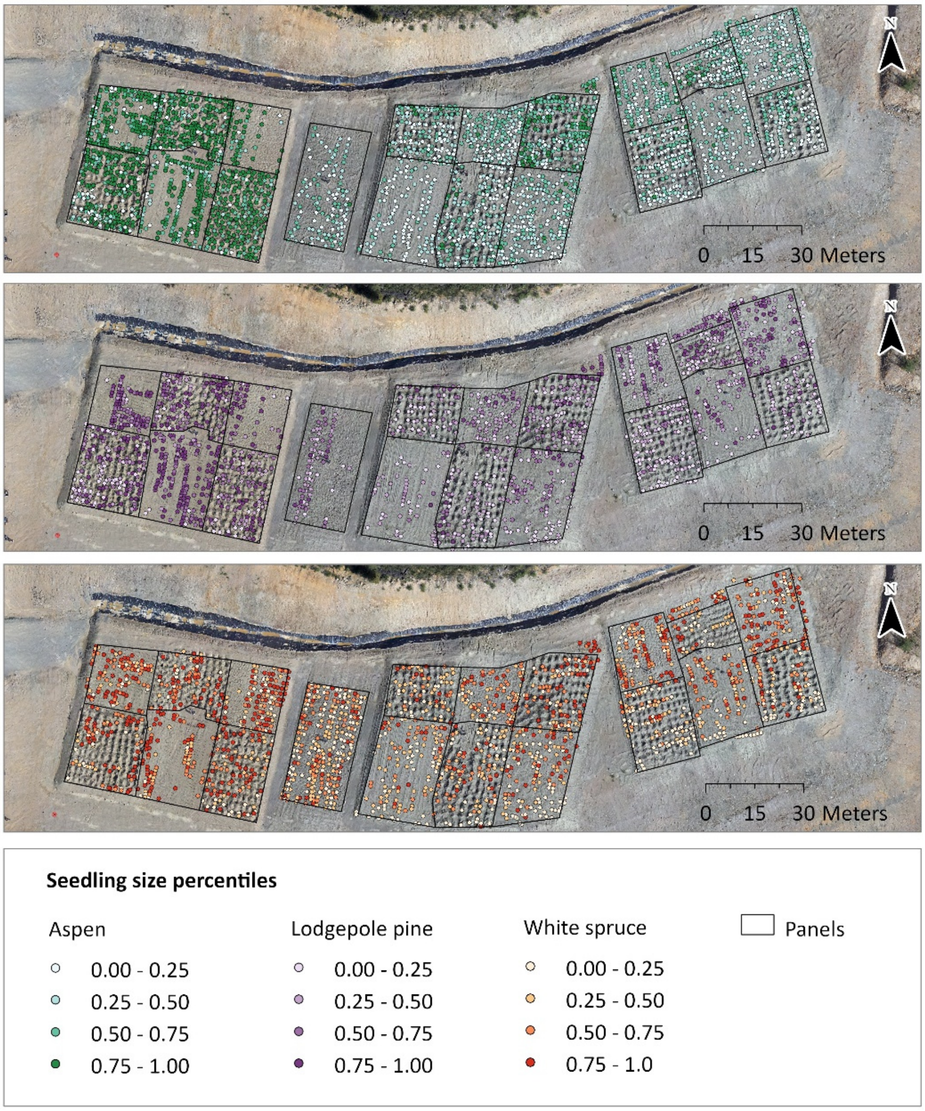

Figure 3 Spatial distribution of aspen, lodgepole pine, and white spruce, classified using a Random Forest model, and seedling size percentiles. Percentiles represent each seedling's size relative to others in the population

4 Discussion

Using a combination of high-spatial-resolution imagery, OBIA, and a Random Forest model, aspen, lodgepole pine, and white spruce seedlings were classified with an average F1-score of 0.88, and seedling size percentiles were shown to be an effective indicator of health. These results suggest that methods developed here can provide scalable, objective, and repeatable methods for tracking seedling health across large areas.

Based on the results of the seedling health analysis, it is recommended that either the Random Forest classifications of survival or seedling size percentiles be used as the primary indicator of seedling health. While the Random Forest classification of seedling vigour showed near-acceptable accuracies, it has some limitations that justify its exclusion. The main issue stems from the subjectivity of field-based vigour estimates and the potential for inter-observer disagreement (Foody 2024). The boundaries between classes, particularly moderate and high, can be ambiguous, leading to inconsistent labelling. With such edge cases, one biologist's moderate vigour call could be another's high. These inconsistencies in the training and validation data can reduce model accuracy and reliability.

Seedling size percentiles generated in this analysis offer a robust, objective, and quantifiable measure for assessing seedling health, avoiding the subjectivity and inconsistency associated with the field-based vigour classifications. Because these percentiles are derived from pixel counts – directly reflecting seedling canopy size – they provide a standardised and repeatable metric.

While the workflow developed in this study produced strong results that support its integration into operational monitoring programs, three key considerations should be addressed: flight altitude, the timing of data collection, and the characteristics of the study area.

A consistent flight altitude relative to the ground must be maintained during the flight. In sloped terrain, if the RPAS maintains a consistent flight altitude above a reference point, rather than following the terrain, pixel sizes at the top and bottom of the slope will differ. However, these pixels will still be treated as if they represent the same ground area, introducing potential error into classification models relying on area as a predictor variable and into the assessment of seedling size based on pixel counts.

For accurate species classification, it is recommended to base the classifications on data collected in the first growing season after planting (after deciduous species have leafed out), when spectral signatures are relatively similar within species. Beyond the first year, within species variation in spectral signatures increases, making species classifications less accurate. This challenge became evident in our attempt to classify species after two summers and one winter, where species-level accuracy was notably reduced. Missing this optimal classification window can limit the reliability of subsequent species-level monitoring, though it should be noted that imagery collected in later years can be useful for manual quality assurance of the initial species classification.

The characteristics of the study area affect the likelihood of successful species classifications and the generation of seedling size percentiles. This study occurred in a newly reclaimed site with little natural ingress, allowing for the clear delineation of individual seedlings. Well-defined boundaries between seedlings are essential for accurately identifying individuals and generating distinct image objects representing each seedling (Feduck et al. 2018), the precursor for both species classifications and the generation of size percentiles. Even after two summers, there was still minimal natural ingress on the plateau, which enabled the isolation of vegetation clusters for monitoring seedling health based on seedling size percentiles. However, these near-bare site conditions are only temporary. Natural ingress will occur over time, eventually obscuring the clean separation between seedlings and surrounding vegetation. This issue was already evident on the slope panels, which were seeded with grasses, where it became more difficult to distinguish seedlings from adjacent cover.

Identifying and delineating seedlings under these conditions is not impossible, but alternative methods must be used, which are generally less effective for very short stature seedlings identified in this study. For example, Jayathunga et al. (2023) utilised 3D point clouds to help filter out non-seedling vegetation. While they were able to identify tree seedlings with high accuracy, seedlings in their study area were mostly 1 m or taller and most seedlings that could not be identified in the point clouds were short with very small crowns. Some seedlings in this study area were as short as 5 cm in height, placing them within the margin of error for all but the most precise remote sensing sensors.

If seedlings are accurately classified and mapped early in the reclamation process, their locations can be used for long-term monitoring, even as spectral separation becomes less reliable. As visual or spectral delineation fades due to vegetation overlap, monitoring strategies can pivot to alternative indicators. For instance, LiDAR-derived height can serve as a robust metric for growth (Castilla et al. 2020), while buffered vegetation indices around known seedling positions could be used to assess changes in vigour (D'Odorico et al. 2020). While these alternative metrics may not capture an increase in horizontal cover of seedlings, they can still provide meaningful insights into seedling performance over time, thereby retaining the value of remote sensing as a monitoring tool as site conditions evolve.

5 Conclusion

This study presents an operational framework for using high-resolution multispectral imagery and machine learning to locate individual seedlings across a study area, classify them into species, and monitor their health over time. By achieving high classification accuracy for aspen, lodgepole pine, and white spruce, and demonstrating the value of seedling size percentiles as a scalable health indicator, this approach offers a robust and objective alternative to traditional field-based monitoring. Future work in this study area will focus on evaluating the impact of different reclamation treatments on seedling development. By analysing species-specific responses to the various treatments applied in this trial, we can provide insights to guide future reclamation efforts, helping to refine treatment strategies, select more effective planting materials, and prioritise species that show strong performance under site conditions.

References

Aasen, H, Honkavaara, E, Lucieer, A & Zarco-Tejada, PJ 2018, 'Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows', *Remote Sensing*, vol. 10, no. 7, p. 1091, https://doi.org/10.3390/rs10071091

Álvarez-Taboada, F, Paredes, C & Julián-Pelaz, J 2017, 'Mapping of the invasive species Hakea sericea using unmanned aerial vehicle (UAV) and WorldView-2 imagery and an object-oriented approach', *Remote Sensing*, vol. 9, no. 9, p. 913, https://doi.org/10.3390/rs9090913

Anderson, J, Iverson, M & Pearse, B 2021, 'On the use of unmanned aerial systems for environmental monitoring', in L Fraser (ed.), *BC Technical Research Committee on Reclamation Symposium*, Vancouver, https://doi.org/10.14288/1.0402661

Asmara, DH, Allaire, S, van Noordwijk, M & Khasa, DP 2022, 'Tree establishment on post-mining waste soils: species, density, and mixture effects', *Canadian Journal of Forest Research*, vol. 52, no. 1, pp. 79–89, https://doi.org/10.1139/cjfr-2021-0044

Bagaram, MB, Giuliarelli, D, Chirici, G, Giannetti, F & Barbati, A 2018, 'UAV remote sensing for biodiversity monitoring: are forest canopy gaps good covariates?', *Remote Sensing*, vol. 10, no. 9, p. 1400, https://doi.org/10.3390/rs10091400

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

Bergquist, J, Örlander, G & Nilsson, U 2003, 'Interactions among forestry regeneration treatments, plant vigour and browsing damage by deer', *New Forests*, vol. 25, no. 1, pp. 1–16, https://doi.org/10.1023/A:1022660417181

Blaschke, T, Hay, GJ, Kelly, M, Lang, S, Hofmann, P, Addink, E, … Tiede, D 2014, 'Geographic object-based image analysis – towards a new paradigm', *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 24–31, https://doi.org/10.1016/j.isprsjprs.2013.09.014

Breiman, L 2001, 'Random forests', *Machine Learning*, vol. 45, no. 1, pp. 5–32, https://doi.org/10.1023/A:1010933404324

Castilla, G, Filiatrault, M, McDermid, GJ & Gartrell, M 2020, 'Estimating individual conifer seedling height using drone-based image point clouds', *Forests*, vol. 11, no. 9, p. 980, https://doi.org/10.3390/f11090980

Daniels, L, Eeckhout, E, Wieme, J, Dejaegher, Y, Audenaert, K & Maes, WH 2023, 'Identifying the optimal radiometric calibration method for UAV-based multispectral imaging', *Remote Sensing*, vol. 15, no. 11, p. 2800, https://doi.org/10.3390/rs15112800

Deshaies, A, Boudreau S & Harper KA 2009, 'Assisted revegetation in a subarctic environment: effects of fertilisation on the performance of three indigenous plant species', *Arctic, Antarctic, and Alpine Research*, vol. 41, no. 4, pp. 523–532, https://doi.org/10.1657/1938-4246-41.4.523

D'Odorico, P, Besik, A, Wong, CYS, Isabel, N & Ensminger, I 2020, 'High-throughput drone-based remote sensing reliably tracks phenology in thousands of conifer seedlings', *New Phytologist*, vol. 226, no. 6, pp. 1891–1904, https://doi.org/10.1111/nph.16400

Fassnacht, FE, Latifi, H, Stereńczak, K, Modzelewska, A, Lefsky, M, Waser, LT, … Ghosh, A 2016, 'Review of studies on tree species classification from remotely sensed data', *Remote Sensing of Environment*, vol. 186, pp. 112–128, https://doi.org/10.1016/j.rse.2016.08.013

Feduck, C, McDermid, GJ & Castilla, G 2018, 'Detection of coniferous seedlings in UAV imagery', *Forests*, vol. 9, no. 7, p. 384, https://doi.org/10.3390/f9070384

Fernández, N, Ferrier, S & Navarro, LM 2020, 'Essential biodiversity variables: integrating in-situ observations and remote sensing through modeling', in HM Pereira (ed.), *Remote Sensing of Plant Biodiversity*, Springer International Publishing, Vancouver, pp. 485–501.

Foody, GM 2024, 'Ground truth in classification accuracy assessment: myth and reality', *Geomatics*, vol. 4, no. 1, pp. 81–90, https://doi.org/10.3390/geomatics4010005

Gillies, S 2019, *Rasterio: Geospatial Raster I/O for Python Programmers,* MapBox, computer software, https://github.com/rasterio/rasterio

Guerra-Hernández, J, Díaz-Varela, RA, Álvarez-González, JG & Rodríguez-González, PM 2021, 'Assessing a novel modelling approach with high-resolution UAV imagery for monitoring health status in priority riparian forests', *Forest Ecosystems*, vol. 8, p. 43, https://doi.org/10.1186/s40663-021-00342-w

Hologa, R, Scheffczyk, K, Dreiser, C & Gärtner, S 2021, 'Tree species classification in a temperate mixed mountain forest landscape using random forest and multiple datasets', *Remote Sensing*, vol. 13, no. 22, p. 4683, https://doi.org/10.3390/rs13224683

Homer, C, Huang, C, Yang, L, Wylie, B & Coan M 2004, 'Development of a 2001 national land-cover database for the United States', *Photogrammetric Engineering & Remote Sensing*, vol. 70, no. 7, pp. 829–840, https://doi.org/10.14358/PERS.70.7.829

Iizuka, K, Itoh, M, Shiodera, S, Matsubara, T, Dohar, M & Watanabe, K 2018, 'Advantages of unmanned aerial vehicle (UAV) photogrammetry for landscape analysis compared with satellite data: a case study of post-mining sites in Indonesia', *Cogent Geoscience*, vol. 4, no. 1, p. 1498180, https://doi.org/10.1080/23312041.2018.1498180

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 Sensing*, vol. 4, no. 9, pp. 2661–2693, https://doi.org/10.3390/rs4092661

Janga, B, Asamani, GP, Sun, Z & Cristea, N 2023, 'A review of practical AI for remote sensing in Earth sciences', *Remote Sensing*, vol. 15, no. 16, p. 4057, https://doi.org/10.3390/rs15164057

Jayathunga, S, Pearse, GD & Watt, MS 2023, 'Unsupervised methodology for large-scale tree seedling mapping in diverse forestry settings using UAV-based RGB imagery', *Remote Sensing*, vol. 15, no. 22, p. 5422, https://doi.org/10.3390/rs15225422

Jordahl, K, Bossche, JV, Fleischmann, M, Wasserman, J, McBride, J, Gerard, J, … Leblanc, F 2020, *geopandas/geopandas: V0.8.1,* computer software, Zenodo, https://doi.org/10.5281/zenodo.3946761

Klosterman, S, Melaas, E, Wang, JA, Martínez, A, Frederick, S, O'Keefe, J … Richardson AD 2018, 'Fine-scale perspectives on landscape phenology from unmanned aerial vehicle photography', *Agricultural and Forest Meteorology*, vol. 248, pp. 424–436, https://doi.org/10.1016/j.agrformet.2017.10.015

Lemaître, G, Nogueira, F & Aridas, CK 2017, 'Imbalanced-learn: a Python toolbox to tackle the curse of imbalanced datasets in machine learning', *Journal of Machine Learning Research*, vol. 18, no. 17, pp. 1–5.

Manfreda, S, McCabe, M, Miller, P, Lucas, R, Madrigal, VP, Mallinis, G, … Szabó, B 2018, 'On the use of unmanned aerial systems for environmental monitoring', *Remote Sensing*, vol. 10, no. 4, https://doi.org/10.3390/rs10040641

Maxwell, AE, Warner, TA & Fang, F 2018, 'Implementation of machine-learning classification in remote sensing: an applied review', *International Journal of Remote Sensing*, vol. 39, no. 9, pp. 2784–2817, https://doi.org/10.1080/01431161.2018.1433343

Meng, D, Bao, N, Tayier, K, Li, Q & Yang, T 2024, 'A remote sensing-based index for assessing long-term ecological impact in arid mined land', *Environmental and Sustainability Indicators*, vol. 22, no. 100364, https://doi.org/10.1016/j.indic.2024.100364

Näsi, R, Honkavaara, E, Lyytikäinen-Saarenmaa, P, Blomqvist, M, Litkey, P, Hakala, T, … Holopainen, M 2015, 'Using UAV-based photogrammetry and hyperspectral imaging for mapping bark beetle damage at tree level', *Remote Sensing*, vol. 7, no. 11, pp. 15467–15493, https://doi.org/10.3390/rs71115467

Pan, W, Wang, X, Sun, Y, Wang, J, Li, Y & Li, S 2023, 'Karst vegetation coverage detection using UAV multispectral vegetation indices and a machine learning algorithm', *Plant Methods*, vol. 19, article 7, https://doi.org/10.1186/s13007-023-00982-7

Polster, DF 2016, 'Natural processes for the restoration of drastically disturbed sites', *Journal of the American Society of Mining and Reclamation*, vol. 2016, no. 2, pp. 77–90, https://doi.org/10.21000/JASMR16020077

Pedregosa, F, Varoquaux, G, Gramfort, A, Michel, V, Thirion, B, Grisel, O, … Duchesnay, É 2018, *Scikit-learn: Machine Learning in Python*, computer software, arXiv, https://doi.org/10.48550/arXiv.1201.0490

Reinprecht, V & Kieffer, DS 2025, 'Application of UAV photogrammetry and multispectral image analysis for identifying land use and vegetation cover succession in former mining areas', *Remote Sensing*, vol. 17, no. 3, p. 405, https://doi.org/10.3390/rs17030405

Saarinen, N, Vastaranta, M, Näsi, R, Rosnell, T, Hakala, T, Honkavaara, E, … Hyyppä, J 2018, 'Assessing biodiversity in boreal forests with UAV-based photogrammetric point clouds and hyperspectral imaging', *Remote Sensing*, vol. 10, no. 2, p. 338, https://doi.org/10.3390/rs10020338

Śledź, S & Ewertowski, MW 2022, 'Evaluation of the influence of processing parameters in structure-from-motion software on the quality of digital elevation models and orthomosaics in the context of studies on earth surface dynamics', *Remote Sensing*, vol. 14, no. 6, p. 1312, https://doi.org/10.3390/rs14061312

Tordoff, GM, Baker, AJM & Willis, AJ 2000, 'Current approaches to the revegetation and reclamation of metalliferous mine wastes', *Chemosphere*, vol. 41, no. 1–2, pp. 219–228, https://doi.org/10.1016/S0045-6535(99)00414-2

Vanegas, F, Bratanov, D, Powell, K, Weiss, J & Gonzalez, F 2018, 'A novel methodology for improving plant pest surveillance in vineyards and crops using UAV-based hyperspectral and spatial data', *Sensors*, vol. 18, no. 1, p. 260, https://doi.org/10.3390/s18010260

Xie, Y, Sha, Z & Yu, M 2008, 'Remote sensing imagery in vegetation mapping: a review', *Journal of Plant Ecology*, vol. 1, no. 1, pp. 9–23, https://doi.org/10.1093/jpe/rtm005