



REMOTE SENSING TOOLS FOR ENVIRONMENTAL MONITORING AND CERTIFICATION OF WELL SITES

Michael Henley, Vanessa Caron, Parnian Rezaia, Eduardo Loos,
Kevin Renkema, and Dean MacKenzie, Vertex Professional Services
Ltd.

REPORT PREPARED FOR

PETROLEUM TECHNOLOGY ALLIANCE CANADA

Reclamation Remediation Research Committee

CONFIDENTIAL

20-RRRC-13

March 31, 2024

DISCLAIMER

PTAC does not warrant or make any representations or claims as to the validity, accuracy, currency, timeliness, completeness or otherwise of the information contained in this report, nor shall it be liable or responsible for any claim or damage, direct, indirect, special, consequential, or otherwise arising out of the interpretation, use or reliance upon, authorized or unauthorized, of such information.

The material and information in this report are being made available only under the conditions set out herein. PTAC reserves rights to the intellectual property presented in this report, which includes, but is not limited to, our copyrights, trademarks, and corporate logos. No material from this report may be copied, reproduced, republished, uploaded, posted, transmitted, or distributed in any way, unless otherwise indicated on this report, except for your own personal or internal company use.

CITATION

This report may be cited as:

Caron, V.; Henley, M.; Rezanian, P.; Loos, E.; Renkema, K.; MacKenzie, D. 2024. Remote Sensing Tools for Environmental Monitoring and Certification of Well Sites. Prepared for the Petroleum Technology Alliance of Canada, Calgary, Alberta. Report 20-RRRC-13. 40 p.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial contributions provided by the Alberta Upstream Research Fund (AUPRF) Program and the Clean Resource Innovation Network (CRIN) as well as the guidance and support provided by the technical project champions (Sonia Glubish and Lisa Warren). We would also like to acknowledge Emily Herdman (InnoTech Alberta) who provided some of the data used in this project.

EXECUTIVE SUMMARY

This document presents a comprehensive analysis of the use of remote sensing tools for environmental monitoring and the certification of well sites. The study addresses the need for advanced methodologies in the reclamation of padded sites, which are complex due to their construction in sensitive areas. Traditionally, the environmental monitoring of such sites is both time-consuming and potentially disruptive. This project aimed to leverage remote sensing technologies to enhance the efficiency and accuracy of monitoring these environments, thereby supporting better decision-making processes in land use and reclamation certification.

We describe a multi-staged approach, starting with the collection and analysis of extensive datasets from well databases, company records, environmental reports, and remote sensing data. Five tasks were identified to achieve our goals:

Task 1 - Integration of Remote Sensing Data with *In Situ* Data

Task 2- Verification and Assessment of Sustainability of Forests Developing on Pads

Task 3 - Development of Digital Remote Sensing Tools to Detect Impacts Caused by Pads on Peatlands

Task 4 - Differentiation of Bogs and Fens and Associated Well Sites

Task 5 - Development And Improvement of Existing Remote Sensing Tools for Non-Padded Well Sites in The Green Zone

The research utilized a combination of airborne and spaceborne datasets to identify padded sites across Alberta. Advanced machine learning algorithms were trained using a dataset of identified padded and unpadded sites to classify new sites based on remote sensing data. The methodology emphasized the integration of these tools into a decision framework that can systematically be used to complement existing evaluation and monitoring methods to assess the status of reclamation efforts.

We highlight the capabilities of remote sensing tools in detecting and classifying padded sites with considerable accuracy including their importance in reducing the ecological footprint of monitoring activities by minimizing ground disturbances during data collection. However, we acknowledge the challenges in the current remote sensing approach, such as the need for larger training datasets and the potential benefits of incorporating additional predictor variables and alternative classification algorithms.

We demonstrate the effectiveness of remote sensing tools in enhancing the reclamation certification process for well sites in general. These tools not only improve the speed and reduce the costs associated with environmental monitoring but also increase the reliability of the data collected, thus facilitating more informed decision-making. Finally, we recommend the continued

development and integration of these technologies into standard monitoring practices, suggesting that they hold significant promise for widespread application in environmental monitoring and site reclamation across various industries.

TABLE OF CONTENTS

DISCLAIMER.....	i
CITATION.....	i
ACKNOWLEDGEMENTS.....	i
EXECUTIVE SUMMARY	ii
TABLE OF CONTENTS.....	IV
LIST OF FIGURES.....	vi
ACRONYMS	viii
1.0 BACKGROUND	1
1.1 PROJECT OVERVIEW	1
1.2 <i>TASK 1 - INTEGRATION OF REMOTE SENSING DATA WITH IN SITU DATA</i>	1
1.3 <i>TASK 2 - VERIFICATION AND ASSESSMENT OF SUSTAINABILITY OF FORESTS DEVELOPING ON PADS..</i>	1
1.4 <i>TASK 3 - DEVELOPMENT OF DIGITAL REMOTE SENSING TOOLS TO DETECT IMPACTS CAUSED BY PADS</i> <i>ON PEATLANDS.....</i>	2
1.5 <i>TASK 4 - DIFFERENTIATION OF BOGS AND FENS AND ASSOCIATED WELL SITES</i>	2
1.6 <i>TASK 5 - DEVELOPMENT AND IMPROVEMENT OF EXISTING REMOTE SENSING TOOLS FOR NON-</i> <i>PADDED WELL SITES IN THE GREEN ZONE.....</i>	2
2.0 DATA	3
2.1 <i>TASK 1 DATA.....</i>	3
2.1.1 In situ Field Data	3
2.2 <i>TASK 2 DATA.....</i>	3
2.2.1 LiDAR.....	3
2.2.2 Optical Imagery	5
2.3 <i>TASK 3 DATA.....</i>	5
2.3.1 Optical Imagery	5
2.3.2 Site Locations.....	5
2.4 <i>TASK 4 DATA.....</i>	5
2.4.1 Optical Imagery	5
2.4.2 SAR Data	5
2.4.3 Wetland Data.....	6
2.5 <i>TASK 5 DATA.....</i>	6
2.5.1 LiDAR Imagery	6
2.5.1 Optical Imagery	6
3.0 METHODOLOGY	6
3.1 TASK 1	6
3.2 TASK 2.....	6
3.2.1 Site Selection	7

3.2.2	Trend Extraction	9
3.3	TASK 3	9
3.3.1	Site Selection	10
3.3.2	Wetness and Vegetation Index Sampling.....	14
3.4	TASK 4	14
3.4.1	Study Area and Training Data	14
3.4.2	Multi-temporal Imagery	16
3.4.3	Classification and Bog/Tamarack mapping	16
3.5	TASK 5	16
3.5.1	LiDAR Processing	17
3.5.2	Disturbance Detection using Time-Series Data	19
4.0	RESULTS.....	19
4.1	TASK 1	19
4.2	TASK 2	21
4.2.1	Tools	21
4.2.2	Outputs.....	21
4.3	TASK 3	25
4.3.1	Tools	25
4.3.2	Outputs.....	25
4.4	TASK 4	29
4.4.1	Tools	29
4.4.2	Outputs.....	29
4.5	TASK 5	31
4.5.1	Tools	32
4.5.2	Outputs.....	32
5.0	DISCUSSION.....	33
5.1	TASK 1	33
5.2	TASK 2	34
5.3	TASK 3	35
5.4	TASK 4	35
5.5	TASK 5	36
6.0	CONCLUSIONS.....	37
7.0	RECOMMENDATIONS	38
7.1	TASK 1	38
7.2	TASK 2	38
7.3	TASK 3	38
7.4	TASK 4	39
7.5	TASK 5	39
8.0	APPLICATION.....	39
9.0	REFERENCES	40

LIST OF FIGURES

Figure 1. The yellow polygon shows the LiDAR coverage provided by Innotech.	4
Figure 2. Location of the 20 study sites (yellow circles) used to for developing the forest sustainability assessment tools.	8
Figure 3. Access roads (red polygons) used in Task 3 for assessing impacts on peatlands.....	11
Figure 4. An example of the automated sampling point creation is shown in this example. The input access roads have a sampling pair of points created at regular intervals and offsets for each feature.	13
Figure 5. Overview of the subset area and training sites used to test the classification tools. Using a smaller subset area greatly simplifies the collection of additional land cover classes and the mosaicking of Sentinel-2 imagery.	15
Figure 6 This figure shows the Bare Earth DEM (left), the Full Feature DEM (centre), and the derived (FFDEM - BEDEM) Canopy Heigh Model (right). The study site polygon is shown in red and the buffered area around the site is shown in yellow.	17
Figure 7. Resampling of LiDAR data at 0.5 m resolution (blue) and the LiDAR resampled to 5 m resolution (green), where each pixel is greater than 2 m in height.	18
Figure 8. Locations of tamarack forest derived from Vertex field data.	20
Figure 9. Report layout of results from the LiDAR-derived sustainability tool. For each site processed, we present three pieces of information: the forest polygons extracted from the 5 m resolution LiDAR data overlain on 50 cm reference imagery, shown in the top frame; the 10 m Sentinel-2 NDVI pixel-based trend map, shown in the bottom frame; and at left is a graph of the NDVI trend for each polygon which is the median value of all pixels within.	23
Figure 10. This figure is another example of the output from the forest sustainability tool, where the results are a more complex in that we have five detected forest areas.	24
Figure 11. Example 1 from Task 3 output. The top frame shows the assessed access road, the greyscale NDVI image from input year (2021), and the classified sampling points along the road. There are four consecutive points that indicate potential impact. In the bottom frame, we can see a wetland area, indicated by the purple arrows, that is the cause for these lower values.	27
Figure 12. Example 2 from Task 3 output. In this example, we can see that several points along the southern side of the feature are classified as potential impacts. The optical imagery in the bottom frame, from the year prior to the assessed year, shows clear evidence of ponding water. Not all ponds are detected by the sampling pairs, see bottom frame, and there are two probably reasons for this: first is that wetness conditions may have been different between the year of assessment (2021) compared to the optical image (2020), and second, an increase of 1 pixel in the offset distance would have resulted in a positive detection.	28
Figure 13. Classification output showing tamarack forest areas detected by the multi-sensor, multi-date supervised classification.	30

Figure 14. Map of tamarack forest classified using our methods, bogs as mapped by the AMWI, and bog/tamarack overlap areas. We can see that there are several large areas of overlap which could indicate potential misclassification of bogs.	31
Figure 15. (A) Sample output of the CHM calculated using from LiDAR at native resolution of 0.5m. (B) Resulting forested areas extracted using a minimum height threshold of 2 m.	32
Figure 16. This figure shows the variability of NDVI throughout a dense time series for a 100 m by 100 m test polygon. The green line shows the annual observed NDVI calculated for the entire polygon. The orange line represents the values based on our implementation of a smoothing algorithm designed to remove atmospheric and other noise.	33

ACRONYMS

ABMI	Alberta Biodiversity Monitoring Institute
AI	Artificial Intelligence
BEDEM	Bare Earth Digital Elevation Model
CHM	Canopy Height Model
DEP	Derived Ecosite Phase
DEM	Digital Elevation Model
DIDs	Alberta Digital Integrated Dispositions
EO	Earth Observation
FFEDM	Full Feature Digital Elevation Model
GEE	Google Earth Engine
LiDAR	Light Detection and Ranging
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
ReMoTo	Reclamation Monitoring Toolbox
SAR	Synthetic Aperture Radar

1.0 BACKGROUND

1.1 PROJECT OVERVIEW

The main goal of this project was to test, validate, and develop digital tools to monitor terrestrial and aquatic environments impacted by oil and gas operations. This was achieved by developing automated cutting-edge Machine Learning (ML) and Artificial Intelligence (AI) tools to process and analyse large volumes of remotely-sensed spaceborne optical and Synthetic Aperture RADAR (SAR) data quickly and efficiently. These tools allow us to optimize the way environmental monitoring is currently conducted through improvements provided by the use and analysis of Earth Observation (EO) imagery and related spatial data.

In contrast to conventional field-based monitoring campaigns, remote sensing analysis of EO imagery can reduce damage to crops and other vegetation because there is no direct contact with the surface, and can thus also reduce the spread of diseases, pathogens, and weeds. The tools developed in this project showcase the ability of EO data and remote sensing techniques to address environmental monitoring of reclamation efforts.

We identified five specific tasks that enable us to achieve our goals, as follows:

1.2 TASK 1 - INTEGRATION OF REMOTE SENSING DATA WITH *IN SITU* DATA

Remote sensing data and *in situ* data were combined to ensure efficient monitoring of terrestrial and wetland environments impacted by oil and gas operations. Remote sensing data provide proxy measurements of key site characteristics, such as vegetation cover, plant health, presence of bare area, pad characteristics, and pooling water, etc. *In situ* observations usually deliver very accurate measurements related to soils, vegetation condition, species composition, ponding water, and local topography. The integration of satellite-based EO data with *in situ* data can reduce uncertainty and improve the reliability of remote sensing terrestrial and wetland datasets.

1.3 TASK 2 - VERIFICATION AND ASSESSMENT OF SUSTAINABILITY OF FORESTS DEVELOPING ON PADS

We explored the ability of existing Light Detection and Ranging (LiDAR) data to identify small, discontinuous areas of forest cover on the padded sites through available canopy height metrics. Time-series of optical data were analyzed in a field data-based ecological recovery model to determine the trajectory of vegetation metrics over those areas that the LiDAR data identify as forested. The combination of these two datasets allowed us to determine whether the forested areas are sustainable, stagnant, or in decline.

1.4 *TASK 3 - DEVELOPMENT OF DIGITAL REMOTE SENSING TOOLS TO DETECT IMPACTS CAUSED BY PADS ON PEATLANDS*

We developed tools that exploit remote sensing techniques to assess the environmental impacts of pads on peatlands. The main advantage of remote sensing is that it provides a wide coverage of areas of interest and high temporal resolution at lower cost when compared to traditional field-based data collection. In addition, we automated image processing and quickly derived terrestrial datasets that can be used to monitor various site characteristics. We explored the potential of ML/AI algorithms to detect significant environmental impacts caused by pads.

1.5 *TASK 4 - DIFFERENTIATION OF BOGS AND FENS AND ASSOCIATED WELL SITES*

Using a combination of spaceborne SAR and optical data, we explored ML/AI algorithms to reliably separate bog peatlands from fen peatlands. This distinction can be important when deciding to remove or leave a padded site in place as its impacts may vary based on the type of wetland in which it is located.

1.6 *TASK 5 - DEVELOPMENT AND IMPROVEMENT OF EXISTING REMOTE SENSING TOOLS FOR NON-PADDED WELL SITES IN THE GREEN ZONE*

We developed and improved our existing tools that exploit remote sensing techniques to assess the environmental impacts of non-padded well sites in the Green Zone. Time-series of hydrological and vegetation data were extracted from optical and SAR data and were analyzed using breakpoint detection algorithms in an ecological disturbance recovery model to detect disturbances and assess long term recovery trends on non-padded well sites.

2.0 DATA

2.1 TASK 1 DATA

2.1.1 In situ Field Data

For our goal of integrating *in situ* field data with remote sensing analysis, we used three different sources of *in situ* field data that contained records of a specific forest cover type, i.e., tamarack, and its location:

- Two separate research programs previously conducted by Vertex for both PTAC and Pathways Alliance where detailed wetland classification and vegetation inventory data had been collected between 2021 and 2023.
- Detailed site assessment data from previous field work from control points where wetlands were classified adjacent to reclaimed oil and gas well sites.

2.2 TASK 2 DATA

2.2.1 LiDAR

LiDAR imagery was supplied by Innotech Alberta and data were collected in 2022 (Figure 1). The data have a horizontal spatial resolution of 1 m and a vertical accuracy of 0.2 m and were provided as Bare Earth Digital Elevation Models (BEDEM) and Full Feature DEM (FFDEM) files in ASCII file format. The point cloud data and hillshade files were provided along with the DEMs, but not used in this project. The data were delivered in 1 km x 1 km tiles.

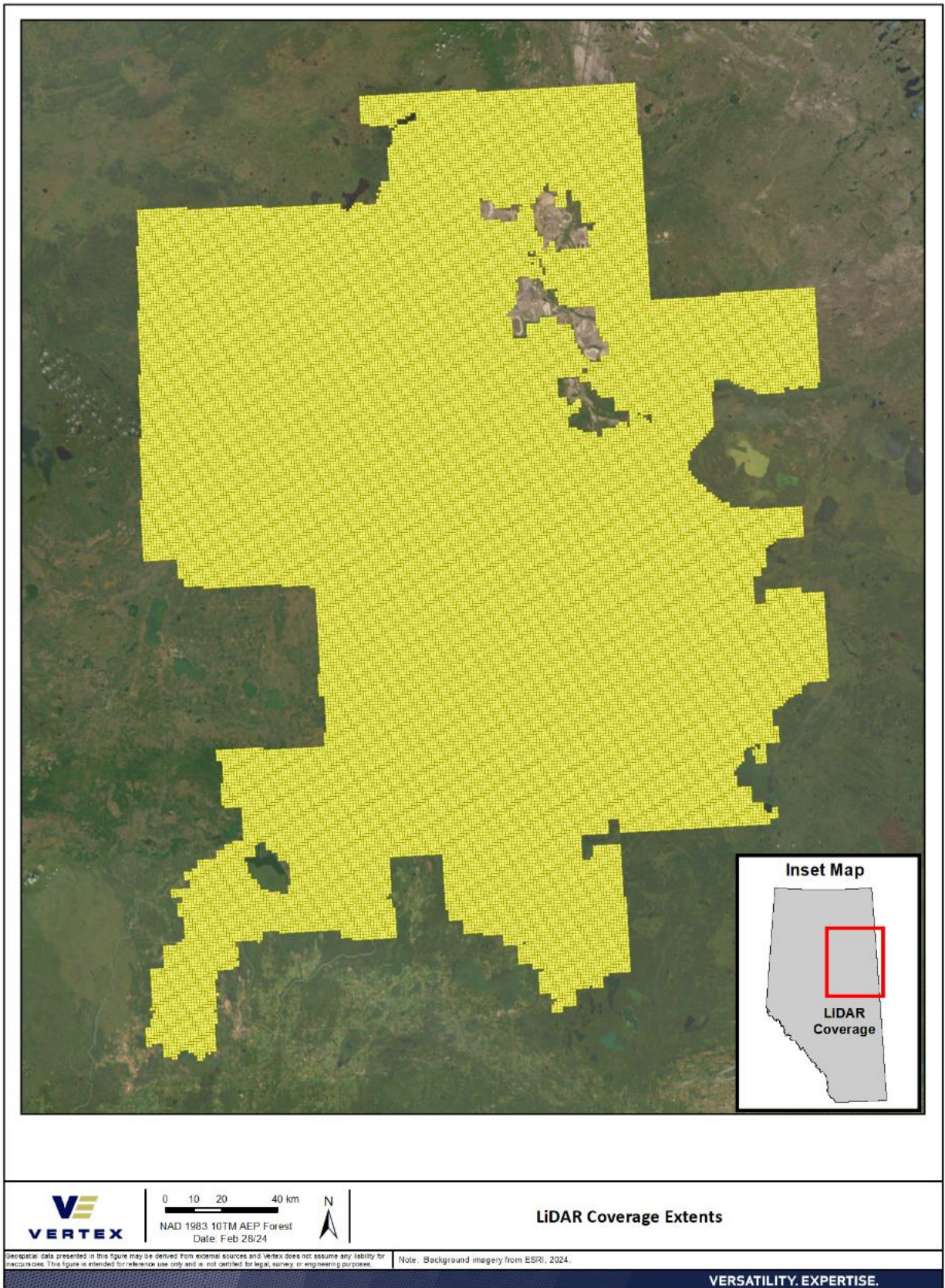


Figure 1. The yellow polygon shows the LiDAR coverage provided by Innotech.

2.2.2 Optical Imagery

Sentinel-2 satellite imagery, supplied by the Copernicus Open Access Hub (European Space Agency, 2022) and accessed via the Google Earth Engine (GEE) platform as 13 optical bands in the visible, near infrared, and short-wave infrared part of the electromagnetic spectrum. The imagery has wavelength-dependent spatial resolution of 10 m, 20 m, and 60 m, that can be used to measure several different characteristics of land cover, such as vegetation or moisture content. Sentinel-2 data were used for this project because (a) they are free and open data, (b) have a high revisit time (5 d), which increases the likelihood of acquiring cloud-free scenes, and (c) their spatial resolution is the finest of all freely-available alternatives like Landsat data (30 m). For this task, we used a time-series of annual best-available-pixel composites from Sentinel imagery between 2017 and 2023.

2.3 TASK 3 DATA

2.3.1 Optical Imagery

Sentinel-2 satellite imagery detailed in Task 2 (Section 2.2.2) were used for this task. We prioritized a single-date 2022 image as the primary data source to allow for meaningful comparisons of along-track landscape conditions.

2.3.2 Site Locations

For the site locations for this task, we selected a set of 106 padded well sites identified in our previous PTAC project (Caron *et al.*, 2022) around Lesser Slave Lake, Alberta. We delineated the access roads for those sites using a combination of Google Earth high spatial resolution imagery and the disposition shapefile data obtained from the Alberta Digital Integrated Dispositions (DIDs) dataset (Alberta Energy Regulator, 2023).

2.4 TASK 4 DATA

2.4.1 Optical Imagery

For this task, we used the cloud-based Sentinel-2 image archive and chose a two-date image pair from the same year to classify tamarack-bearing wetlands. The fall image was acquired on October 09, 2023, and the summer image is from August 20th.

2.4.2 SAR Data

Sentinel-1 calibrated, ortho-corrected imagery collected on October 15, 2023, was used as an additional layer for surface texture information. These data have a resolution of 5 m x 20 m and were collected using a dual polarization. The data are made available by the Copernicus Open Access Hub (European Space Agency, 2022) and accessed via the GEE.

2.4.3 *Wetland Data*

Wetland data obtained from the Alberta Biodiversity Monitoring Institute (ABMI) Wetland Inventory (Alberta Biodiversity Monitoring Institute, 2021) were used to extract a set of wetland polygons that is classified using the training data and optical imagery. These polygons are the overlap points between bog or fen wetlands and the tamarack training data.

2.5 TASK 5 DATA

2.5.1 *LiDAR Imagery*

The first component of this task was to extend our existing tools to unpadded sites and involved adapting our LiDAR processing tools for detecting forested areas on unpadded sites. We used the LiDAR data described in section 2.2.1 to develop this tool.

2.5.1 *Optical Imagery*

The second component of this task uses the Sentinel archive described in 2.2.2.

3.0 METHODOLOGY

3.1 TASK 1

The Vertex field data records described in Section 2.1.1 were used to establish a starting point of known tamarack forest locations. A visual interpretation of high-resolution imagery available in GEE and in ESRI GIS mapping software was used to create a training data point that represents an approximately 10 m² of homogeneous tamarack forest. Care was taken to minimize the inclusion of other land cover types that presented different spectral signatures. These points were collected as a shapefile and were integrated into a larger training dataset, described in Task 4, that were used to classify optical and SAR imagery.

3.2 TASK 2

To assess the sustainability of forests developing on padded sites, we developed an algorithm to determine the vegetation greenness trends of detected forests to describe whether those treed areas are showing trends of increasing, stable, or decreasing greenness. We then compared the observed greenness to surrounding undisturbed treed areas. The output of this tool is a series of polygons outlining forested land cover at each site and trend statistics for those forested areas.

To accomplish this and as part of Task 5, Vertex's existing remote sensing Reclamation Monitoring Tool (ReMoTo) platform takes an input list of study sites and corresponding LiDAR imagery and extracts areas where canopy cover is greater than a specified threshold. It then exports those areas as polygons and leverages cloud computing and available remote sensing

imagery to calculate vegetation greenness metrics over multiple years. These annual measures of vegetation are then analysed to determine the trends of the forest cover polygons and produce polygon and pixel-level forest trend information. This set of LiDAR tools is applicable to both padded and unpadded sites in the Green Zone.

3.2.1 *Site Selection*

The LiDAR coverage we received from Innotech was further east than originally planned for at the outset of the project and did not cover our 2023 field sites that were intended for use as part of Task 1. Given this, we selected 20 study sites (Figure 2) from our previous PTAC project (Caron *et al.*, 2022) that identified padded well sites, and which were within the footprint of the available LiDAR inventory. These sites have a range of tree cover from 0% to approximately 30% that is visible in high spatial resolution imagery. For each study site, three offsite recovery target points were collected to represent undisturbed forest vegetation conditions to which the recovery of the onsite forest conditions were compared.

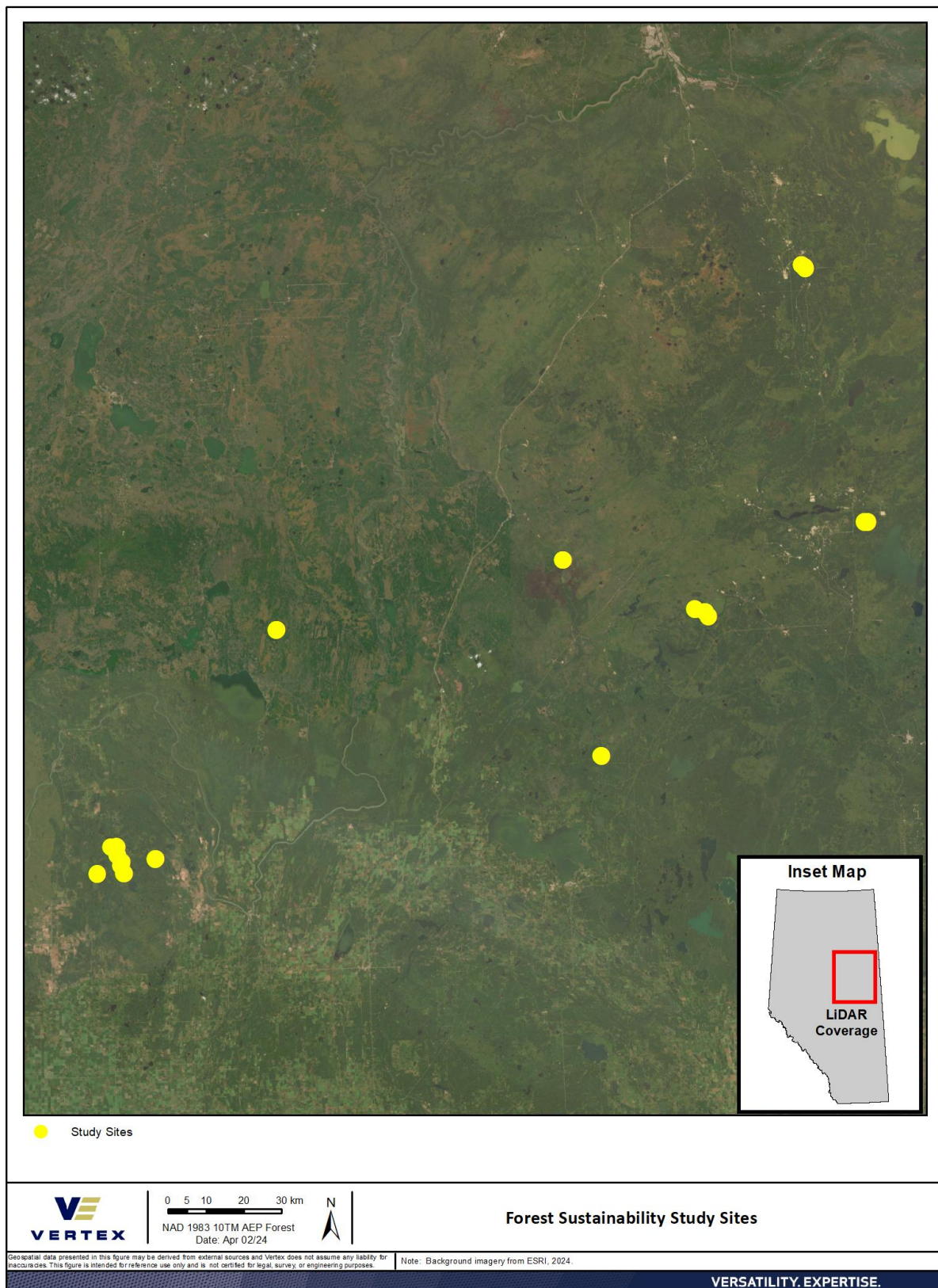


Figure 2. Location of the 20 study sites (yellow circles) used to for developing the forest sustainability assessment tools.

3.2.2 Trend Extraction

Using the newly created forest polygon file from Task 5 that corresponds to trees over 2 m in height, the next step in our processing chain was to create a time-series image stack that can provide the trend of vegetation for each forest polygon. With the forest polygon files uploaded to our ReMoTo platform, we can access both the optical image archives and leverage the power of cloud computing to perform the trend analysis.

ReMoTo creates a time-series image stack used to determine vegetation trends, in one of two ways, specified at the beginning of the processing chain. It can either create a stack between two specified ranges or it can dynamically adjust the length of the time-series. The Sentinel-2 optical data used in this task have a relatively short archive with consistent coverage beginning in 2017. The test sites in this task were all disturbed prior to this date, so data between 2017 and 2023 were used for calculating the trend. Alternately, if using a longer image archive like Landsat, and if each polygon for a project site had a known date of disturbance, we could add a “disturbed date” attribute to our forested area polygons that could be used with the Landsat image archive. This allows ReMoTo to create a site-specific time-series for any disturbed date occurring between 1984 and present day.

With a specified time-series date range input into ReMoTo, the algorithm then calculates a commonly used vegetation index called the Normalized Difference Vegetation Index (NDVI), which is a measure of vegetation greenness. The Sentinel-2 optical image archive for our study area contains seven years of data, from 2017 to 2023. NDVI results were derived annually for every available image within the growing season of that particular year and stacked. The algorithm then searches through this list of available NDVI values at each pixel location and writes out the highest NDVI value found into a new image. This new image is a composite image of the annual best-available-pixel measure of vegetation greenness.

Each of these annual best-pixel images are then stacked into a final time-series dataset for analysis. The final step in the process estimates vegetation recovery trend within the derived forested areas on site by performing a pixel-wise linear regression analysis using a least-square algorithm. The output is a raster image with a trend value for each pixel in the forest polygons.

3.3 TASK 3

In Task 3, our goal was to assess impacts of padded sites on peatlands using EO data. In addition to the padded well sites impacting peatland, the access roads associated with those sites also have potentially measurable impacts on the peatland areas that they pass through. To augment our existing ReMoTo tools that determine whether a well site is padded or unpadded, we developed a new set of tools that assesses the wetness and vegetation conditions along an access road (or any other type of linear feature) and compares the conditions along one side of the road to those on the other side. If the access road is impacting the peatland, there is potentially a

measurable change in the wetness conditions where water ponding can be detected, or significant vegetation change can be mapped.

The tools developed for this task require an input set of linear features. From this set of features, ReMoTo creates a set of sampling points, it calculates wetness and vegetation metrics for each linear feature and samples the wetness and vegetation values incrementally along the linear features. The output is a set of sampling points in shapefile format that are classified into two categories, one for potential impacts, and one for no impact.

3.3.1 *Site Selection*

We created a set of test features from the padded sites we mapped in our previous PTAC project (Caron, *et al.*, 2022). A set of 106 padded sites around Lesser Slave Lake were identified, and the associated access roads were mapped using a combination of high spatial resolution imagery and the DIDs layer (Figure 3).

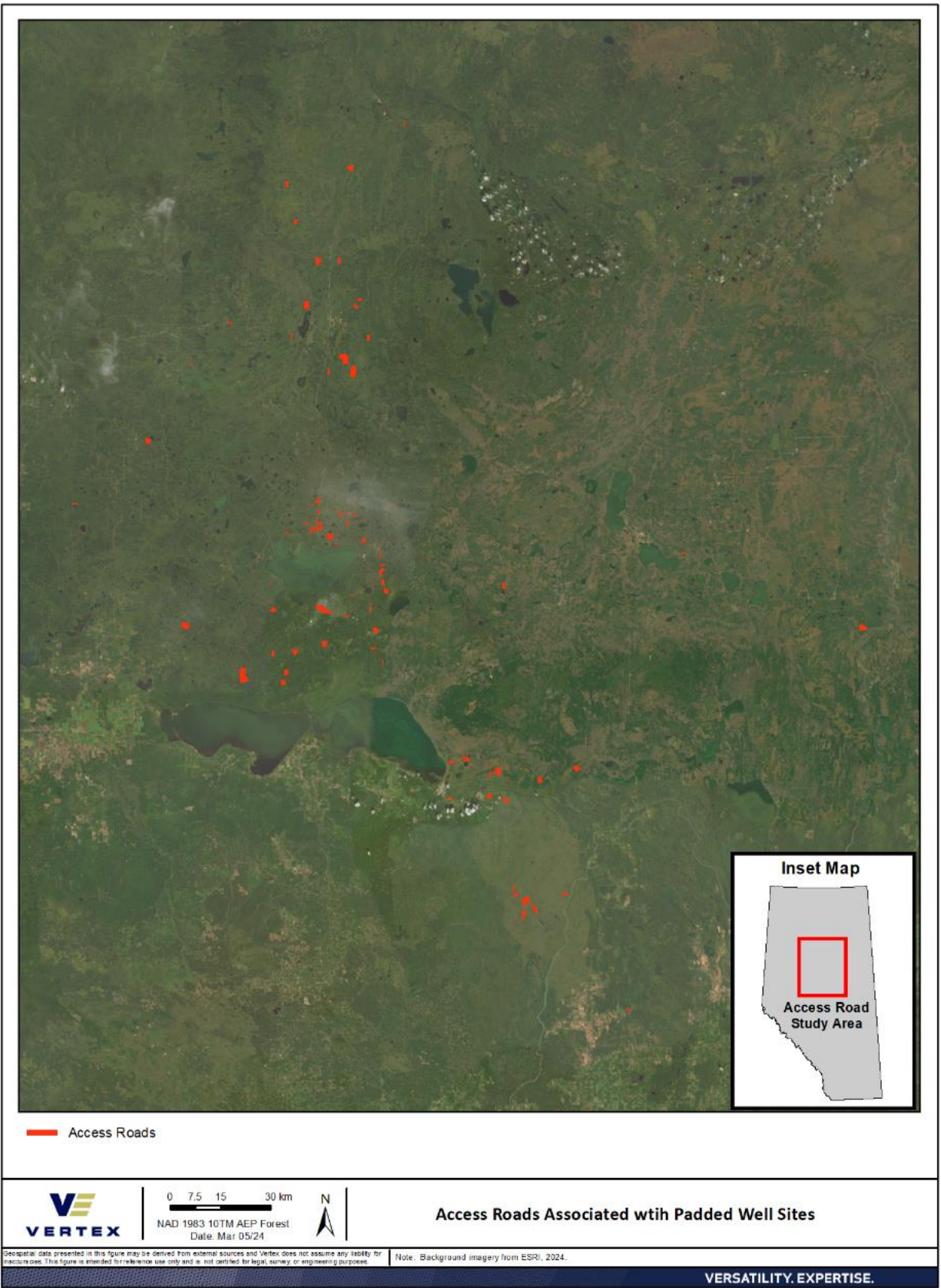


Figure 3. Access roads (red polygons) used in Task 3 for assessing impacts on peatlands.

The roads varied in length from between 50 m and 5 km and pass through variable peatland and non-peatland landscapes, with approximately 81 km in cumulative length. Using this set of roads, ReMoTo then created a set of sampling points at specified intervals along the road, and at specified offsets on either side of the road. The interval and offset are dependent on the spatial resolution of the imagery being used to calculate wetness and vegetation metrics. A short-offset distance between the sampling locations and the road or linear feature was used to increase the likelihood of detecting disturbance-related changes. For this task, using Sentinel-2 optical imagery with a spatial resolution of 10 m, we created sampling points at 50 m intervals along the linear features, and at 15 m offsets left and right of the feature (Figure 4) resulting in 3,120 total sampling locations. ReMoTo created a unique sample ID for every point, and the pairs of points at each interval are also uniquely identifiable so that left and right pairwise comparisons can be made.

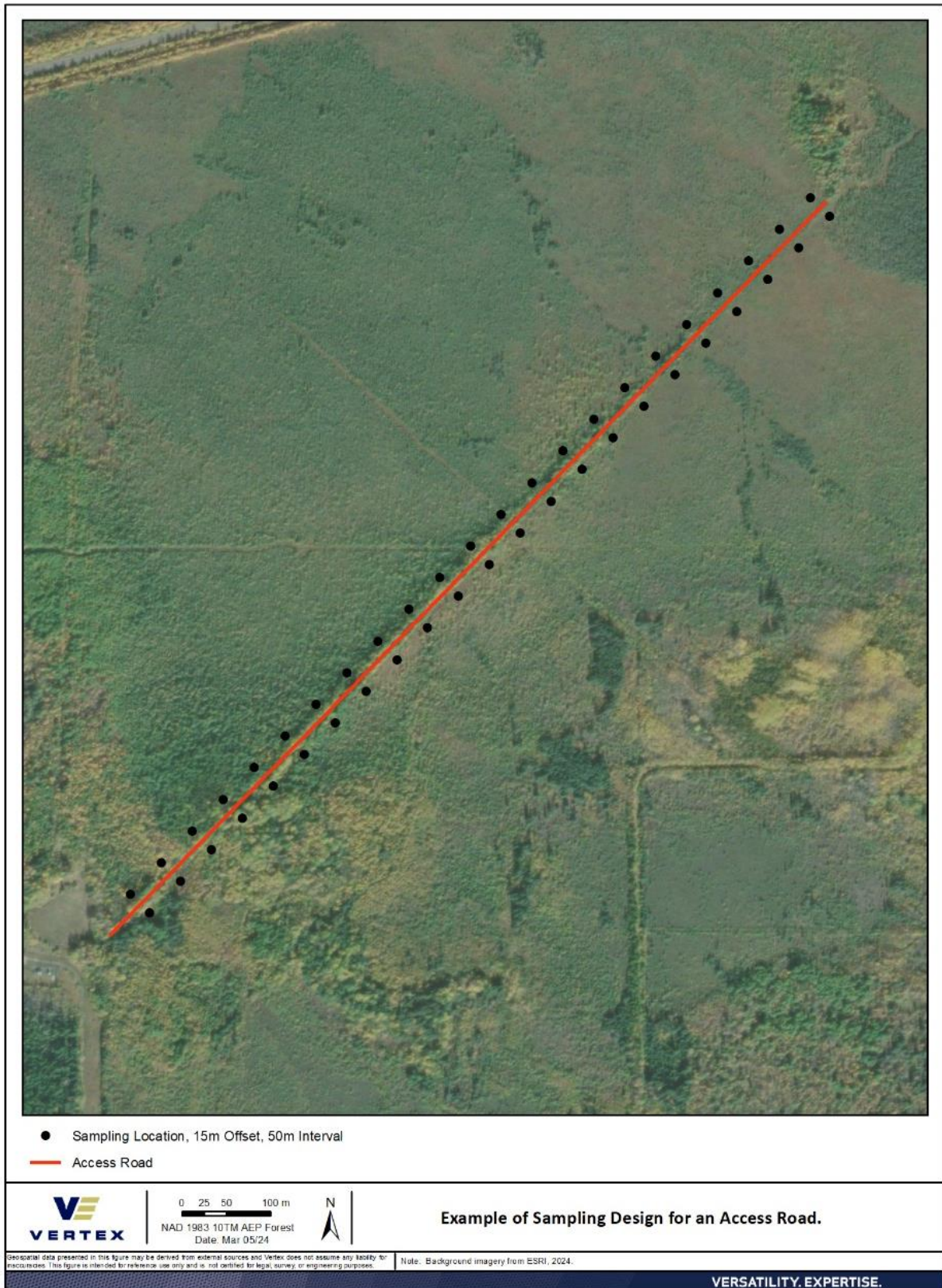


Figure 4. An example of the automated sampling point creation is shown in this example. The input access roads have a sampling pair of points created at regular intervals and offsets for each feature.

3.3.2 *Wetness and Vegetation Index Sampling*

Next, the tool extracts an image from the cloud-based Sentinel archive using a target year specified at the outset of the process. Each sampling point has a wetness metric (Normalized Difference Wetness Index (NDWI)) and a vegetation metric (NDVI) calculated and stored as an attribute. At each interval along the road, the left and right pairs of sampling points are compared. The tool includes a variable that defines a difference tolerance, as a percentage, between the pair of sampling points - if this tolerance is exceeded, the lower of the two values is flagged as a potential impact to be investigated. The final output of the tool is a set of classified sampling points that highlights any areas of significant wetness or vegetation discrepancy.

3.4 TASK 4

We explored ML/AI algorithms to improve the separability of bogs from fens at a finer scale than is generally available with coarse spatial resolution wetland datasets such as the one created by the ABMI or the Derived Ecosite Phase (DEP) Training Data from the Government of Alberta.

There is evidence to suggest that well-developed wetland tamarack forest cover mainly occurs in fen or swamp wetland types (Government of Alberta, 2015). Additionally, our own internal experience conducting field assessments suggests similar distribution and, while tamarack may occur in bogs as individual trees, it does not reach the density that would be observable at the resolution of Sentinel-2 optical imagery. Our field observations do suggest however that fens support larger forest stands of tamarack that may be visible using Sentinel-2 imagery. We tested the ability of Sentinel-2 imagery to detect the unique spectral signature that tamarack trees exhibit during fall senescence.

Using multi-temporal optical and SAR data, in conjunction with the training data created in Task 1, we created a land cover map of tamarack forest cover. The tamarack land cover map was assessed against the existing bog classification available in the ABMI and DEP map data to identify wetlands that may be misclassified as bog. Once identified, further validation using traditional field could be conducted to confirm the results of the analysis.

3.4.1 *Study Area and Training Data*

In Task 1, we used the existing *in situ* data Vertex has acquired over numerous years of field work to create an initial set of tamarack training data. These data are distributed over an area greater

than 20,000 km². To simplify the implementation and testing of the tool, we selected a sub-region of the training data of approximately 200 km² (Figure 5).

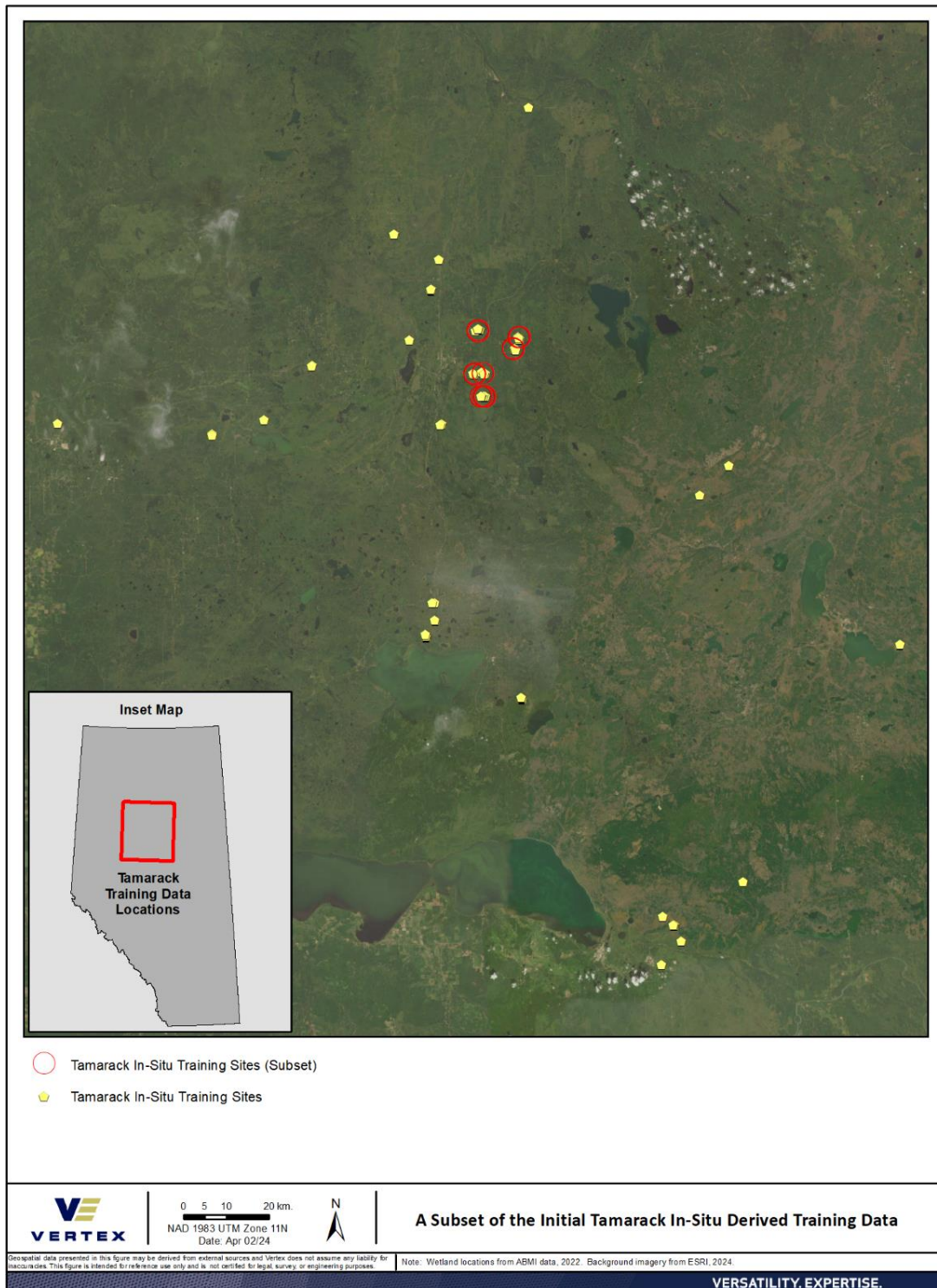


Figure 5. Overview of the subset area and training sites used to test the classification tools. Using a smaller subset area greatly simplifies the collection of additional land cover classes and the mosaicking of Sentinel-2 imagery.

Using the focused study area, we increased the 37 tamarack training sites with additional land cover classes using visual interpretation of high spatial resolution optical imagery. These additional land cover classes included Forest, Shrub and Grass, and Water, and contained between 30 and 60 training locations, depending on the class.

3.4.2 Multi-temporal Imagery

These training data sites were used to classify a multi-temporal Sentinel-2 image stack. ReMoTo was expanded during this project to include the ability to create time-series data stacks that can be as simple as a two-date combination of imagery or include any number of available images from the online data collections. The creation of the multi-temporal image stacks runs in a cloud-processing environment with no local data storage or processing required and has been developed to include calculation and stacking of derived spectral indices. The unique spectral signature exhibited by tamarack in the fall could potentially be confused with other land cover types. To mitigate this, our algorithm uses a two-date image stack of a summer-peak growing season image as well as a fall senescence image. Any potential areas with overlapping spectral signature that are persistent between the two images will have a different cumulative spectral response than tamarack forests where the spectral response progresses from green in the summer to yellow in fall.

3.4.3 Classification and Bog/Tamarack mapping

ReMoTo requires only a few classification parameters to be set up and that we upload our training data to a processing folder in our cloud environment. The output of the classification is a full land cover map of the input training data where each pixel is assigned to a class.

We then converted the raster output to a polygon shapefile and created a new layer that only contains tamarack polygons. We overlaid the map of tamarack forest areas with the existing ABMI bog wetland polygons to create a new dataset that shows existing bog polygons which contain tamarack forest cover greater than 10 m by 10 m in size and potentially misclassified as bogs.

3.5 TASK 5

Work completed for an earlier PTAC project (Caron, et al., 2022) allowed Vertex to develop a method to distinguish padded sites in peatland from unpadded sites, using LiDAR and optical imagery. Extending those tools to applications beyond padded sites involves using those LiDAR tools and adding the ability to create a Canopy Height Model (CHM) that allows us to extract forest polygons used in Task 2. Additionally, software we developed in Task 2 built time-series datasets for annual vegetation trend analysis. Those tools were extended during this task to create a dense time-series tool that uses every available image for a given date range and can identify disturbances that occur infrequently or have only a short observable duration, such as localized flooding, and which may impact long-term recovery.

3.5.1 LiDAR Processing

Our existing tools for LiDAR data processing within ReMoTo include the ability to search for relevant LiDAR tiles from a large data catalogue, extract the elevation datasets, and perform a ML-based classification using those data. Here, we added the ability to create and analyse canopy height to detect forested areas on padded and unpadded sites.

We begin by utilizing a supplied shapefile from the LiDAR delivery that is an index of all available LiDAR data and where each tile is has a unique identifier. We intersected the study site polygons from Task 2 with this Tile ID shapefile and added it to our site list. From there, our software loops through each site independently and extracts both the BEDEM and FFDEM from the matching folder.

Any files in the LiDAR delivery that are not in GeoTiff format are automatically converted and the extraction of forest begins. Our study site polygons are buffered by 100 m to create a larger footprint that will provide context within the surrounding landscape in the reporting stage of the tool. This buffered file is used to create subsets of the BE and FF imagery so we can reduce data processing volumes in subsequent steps.

Next, we subtract the BE pixel value from the FF value for the chip and this results in the CHM, which is a raster image that represents the height of whatever cover is present in the image relative to the ground (Figure 6). For example, if an input FF pixel contains a value of 15 m and the corresponding BE value is 10 m, then the resultant CHM pixel will have a value of 5 m.



Figure 6 This figure shows the Bare Earth DEM (left), the Full Feature DEM (centre), and the derived (FFDEM - BEDEM) Canopy Height Model (right). The study site polygon is shown in red and the buffered area around the site is shown in yellow.

The discrepancy in the spatial resolutions of the Sentinel-2 (10 m) and LiDAR (0.5 m) datasets can result in several small LiDAR polygons that are each within the bounds of a single optical pixel. This increases processing time, complexity, and yields no practical benefit as the same trend information will apply to each high-resolution forest polygon. We mitigated this issue by

resampling the LiDAR to a more directly comparable resolution so that we can extract useful time-series trends from the much coarser optical data. We experimented with a series of resampling values ranging from 1.0 m to 5.0 m – we ended up choosing 5.0 m as having the best compromise between preserving LiDAR information and minimizing redundant information being extracted from the optical data (Figure 7). This resampling value was tailored for the 10 m resolution of Sentinel-2 optical imagery and would need to be adjusted for use with a different time-series data source such as Landsat.

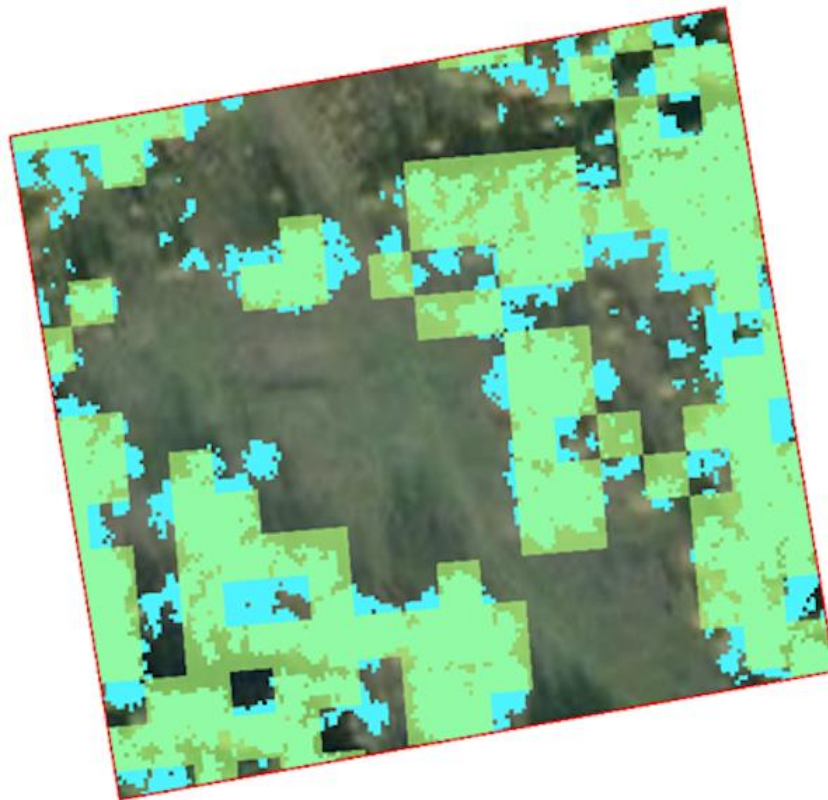


Figure 7. Resampling of LiDAR data at 0.5 m resolution (blue) and the LiDAR resampled to 5 m resolution (green), where each pixel is greater than 2 m in height.

Once the data had been resampled, our final LiDAR processing step was to create a threshold map of all pixels that are greater than an expected minimum height for a forested area. For our testing purposes, we used a value of 2 m as the minimum height for the CHM as shown in the above figure. The green pixels show the LiDAR data resampled to 5 m resolution and they preserve the majority of the original 0.5 m resolution pixels, shown in blue. This resampled raster was then converted to a polygon for use in extracting the historical trend information for each pixel in the study area. At this point, we removed any polygons with an area smaller than two resampled pixels – 50 m². This eliminates any polygons where no more than half of a Sentinel-2 pixel trend information would be derived from non-forest vegetation.

3.5.2 Disturbance Detection using Time-Series Data

In Task 2, we developed software tools to build time-series datasets that can be used to extract trend information from various types of EO data. The tools create time-series datasets in two different ways. The first way is to create a stack of single-date images using a target date of acquisition. For example, if we select August 15 as a peak growing season image, it will cycle through all available imagery and select a cloud-free image from each year that is closest to this date, for all selected years. The second way, and which was implemented in Task 2, is to create a composite best-pixel image for analysing spectral indices. Annual vegetation growth rates vary year by year, so to compensate for this we calculate the NDVI value for all available images in a year or growing season, then choose the highest value that occurs at each pixel in that set of images. This value is then written to a new image that consists of the highest value in that year for each pixel regardless of when it occurred. This is useful for assessing vegetation health or growth but can also hide short-duration changes that may be of interest.

Short-duration impacts, such as localised flooding due to a rain event, are not readily identified using a best-pixel time-series dataset that seeks to find a single highest available pixel value for a given location. To look for these short-duration impacts, we adapted the newly-developed code into a tool that assembles a dense time-series of all available imagery, rather than selected images or composites. For each year in the times-series data range that we specify, all available images are stacked into a single large data cube that contains dozens to hundreds of layers.

Using every available image to create a time-series can give greater insight into localised, short-duration events, but such data are inherently noisy due to inadequate cloud masking, data anomalies, haze, smoke, or other atmospheric effects that can impact data capture. To mitigate this noise, the process includes an option to apply a data smoothing moving-window algorithm based on the median value of the pixel. The output of this tool is a data plot of all spectral index values for the entire date range of interest.

4.0 RESULTS

4.1 Task 1

Using the existing field data as guide to previously known locations of tamarack cover, we positioned 113 training data points within larger areas of uniform tamarack forest cover using high spatial resolution imagery visible in Google Earth and through available data in ESRI mapping software. The training data points that identify tamarack forest cover (Figure 8) that are approximately 10 m in diameter and represent homogeneous forest cover.

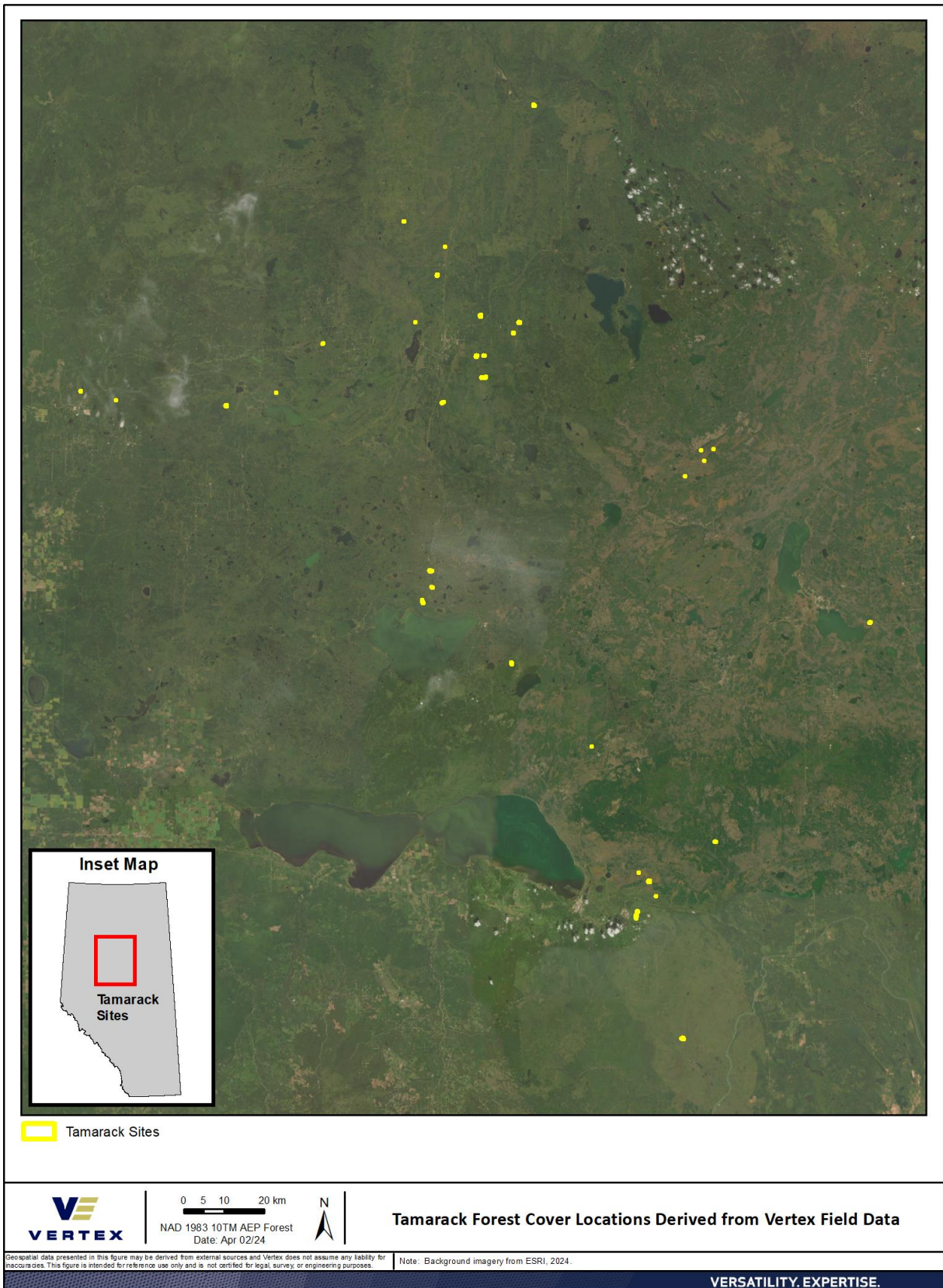


Figure 8. Locations of tamarack forest derived from Vertex field data.

4.2 TASK 2

4.2.1 Tools

The toolset to detect and extract vegetation recovery metrics was successfully developed and tested using a dataset of 20 test sites and corresponding LiDAR imagery. The tools and functionality developed for this task include:

- Eight new modules and functions related to LiDAR processing were developed, tested, and implemented. These include the ability transform LiDAR data formats, calculate CHMs, resample data, create threshold-based feature extraction, and export polygons of treed areas.
- Four new modules were created for the creation of time-series trend data. These include routines that create composite images of best-available pixel indices, create image stacks, employ moving-average smoothing algorithms, calculate linear regression of pixel trends, and export spectral trajectory plots.

4.2.2 Outputs

For every site analysed with our forest sustainability tool, we created a one-page report (Figure 9 and Figure 10) that presents three separate areas of information. The top frame in the report is a map of the forest polygons we extracted from the resampled 5-m resolution LiDAR data and which is overlain on the available high spatial resolution imagery from ESRI. It allows us to inspect the accuracy of the extracted forest polygons based on available imagery, though care must be taken as the available imagery from ESRI is not necessarily from the same year of the LiDAR dataset. Each polygon of forested area is shown in a different colour, and its unique name is shown in the legend to the right of the frame.

The bottom frame is a map of the 10-m Sentinel-2 satellite imagery-based NDVI trends. Each pixel represents the overall slope or trend of the NDVI value in that location over the entire time-series. At the bottom right of the example is a slope legend called “Rate of Annual NDVI Change”. This bottom frame also includes the forest polygons for reference. In this frame, we can see the individual trends of each pixel and look for potential areas of concern.

The “Spectral Trajectory Plot” is the final piece of information included and presents the NDVI trendline for each polygon detected at the site shown in a graph on the right. For each year in the time-series, we calculate the median NDVI value for the entire polygon and plot this over the range of years assessed. This allows us to view the annual trend over time for each forested stand and assess its sustainability as a unit rather than by individual pixels within. Included in this plot is a reference line that represents the median NDVI trends calculated from the surrounding reference plots located in undisturbed forest cover.

In the Figure 9 report, both polygons show stable or slightly increased NDVI trends, with only two pixels showing a slight decline in the purple polygons. Both trendline plots show very shallow increasing values indicating these areas are likely stable and when compared to the reference value averaged from the three reference sample points, the forested areas are at or above the reference value.

The algorithm detected five discrete forested areas, though three of those were quite small (Figure 10). Those smaller forested polygons may overestimate the forest cover present based on the high spatial resolution optical imagery. Further investigation into the validity of these polygons is warranted and the minimum polygon size may need to be altered to increase polygon accuracy.

Within the polygons, we see several large areas of declining NDVI trends evidenced by the light red pixels. These are areas of potential concern for sustainability and are further evidenced by the trendline graphs. While each trendline is roughly at the same value as the reference point average, we can see the overall trend for polygons 1 and 2 show potential decline between 2020 and 2023. This should be interpreted with caution however, due to the very short duration of the time-series, but they represent areas for longer term monitoring.

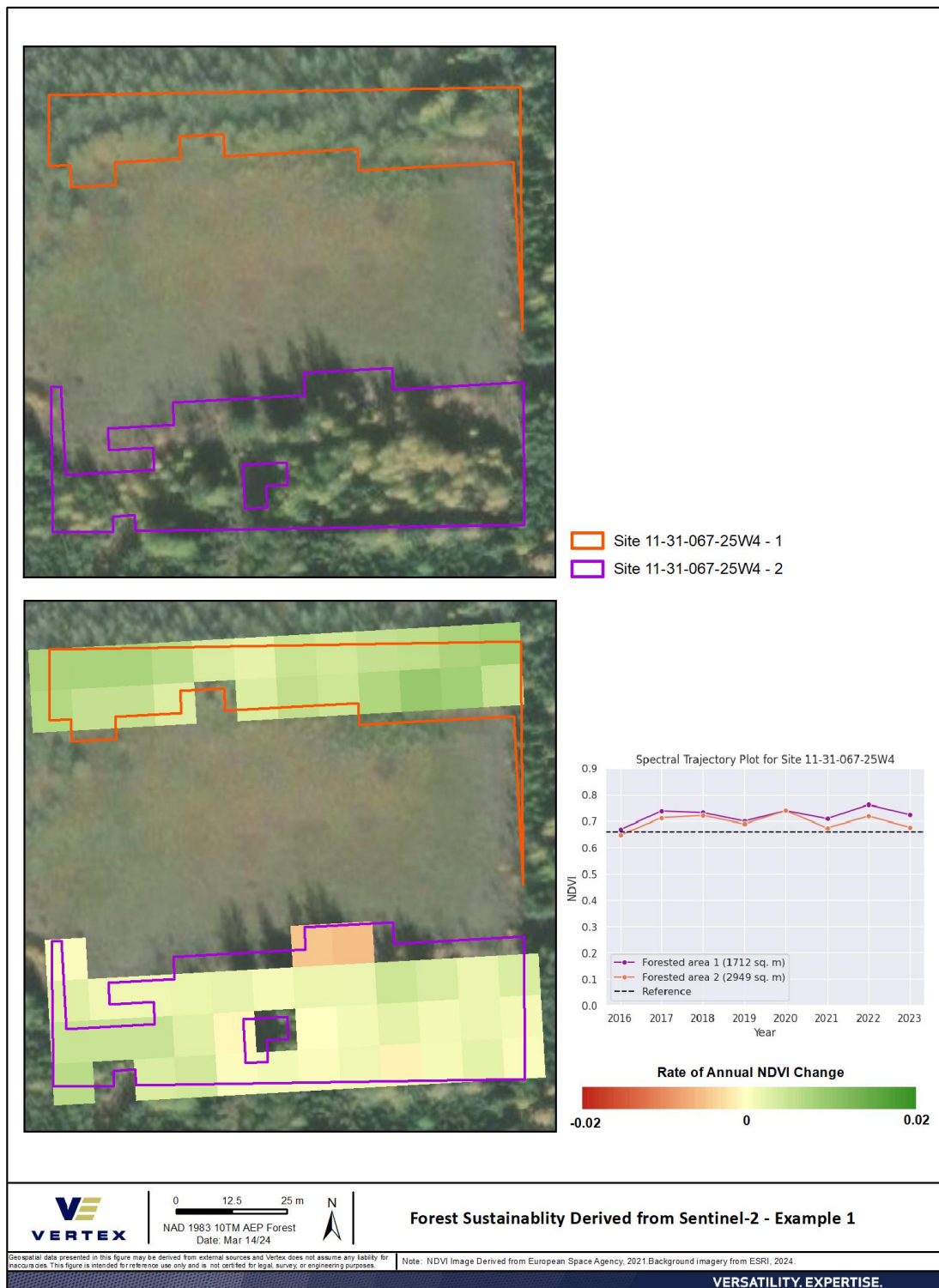


Figure 9. Report layout of results from the LiDAR-derived sustainability tool. For each site processed, we present three pieces of information: the forest polygons extracted from the 5 m resolution LiDAR data overlain on 50 cm reference imagery, shown in the top frame; the 10 m Sentinel-2 NDVI pixel-based trend map, shown in the bottom frame; and at left is a graph of the NDVI trend for each polygon which is the median value of all pixels within.

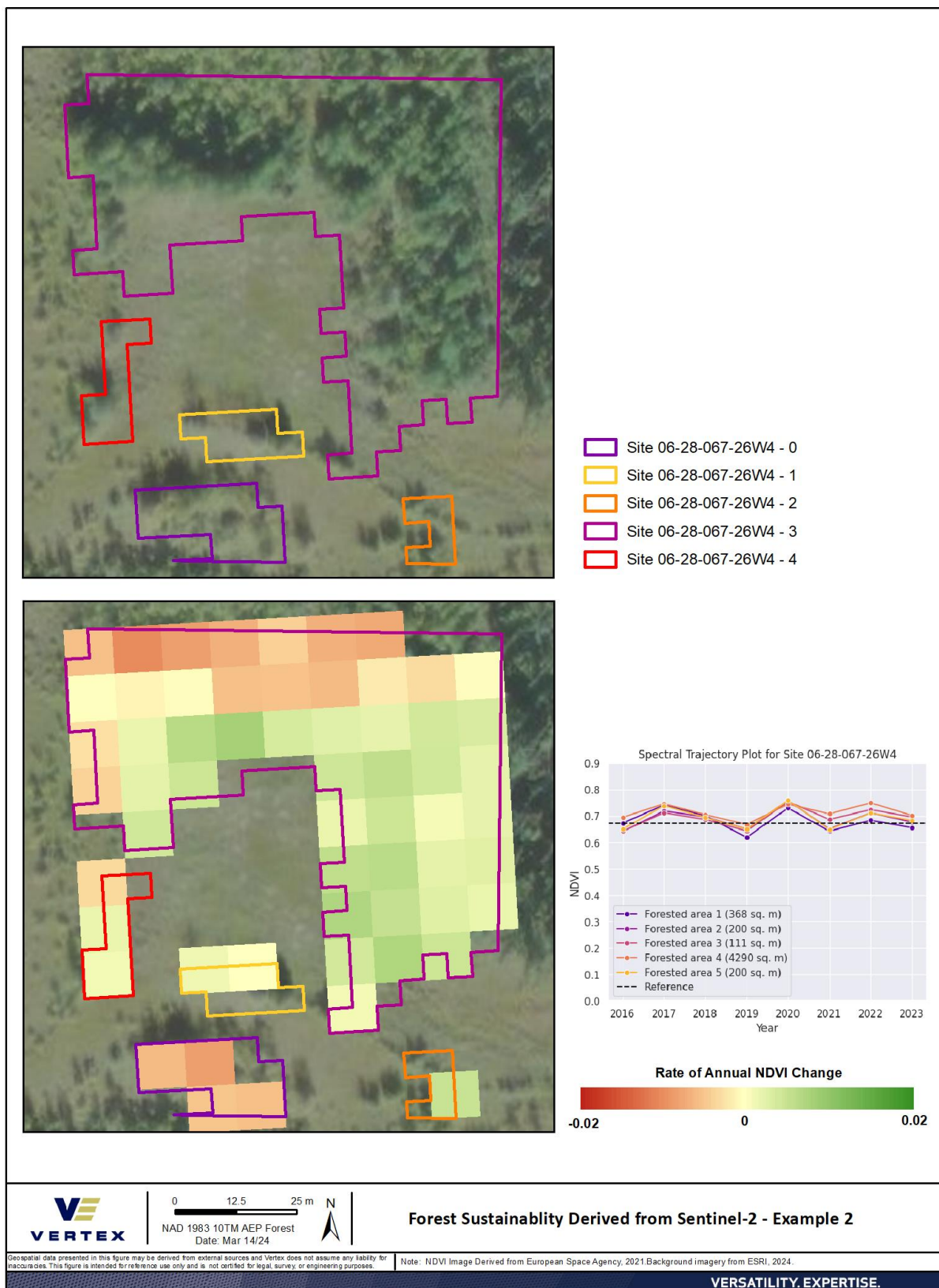


Figure 10. This figure is another example of the output from the forest sustainability tool, where the results are a more complex in that we have five detected forest areas.

4.3 TASK 3

Padded linear features associated with padded well sites may impact peatlands by altering normal hydrological processes that can subsequently impact vegetation health and composition. Methods developed in this Task were designed to detect changes in vegetation and/or water metrics derived from remote sensing data that occur along those linear features.

4.3.1 *Tools*

Several new modules were developed for image analysis that improve the creation of sampling points at data-driven interval and offset distances, creating spectral indices for linear feature sites, extracting spectral index values based on sampling pairs, and assigning those values to potential impact categories.

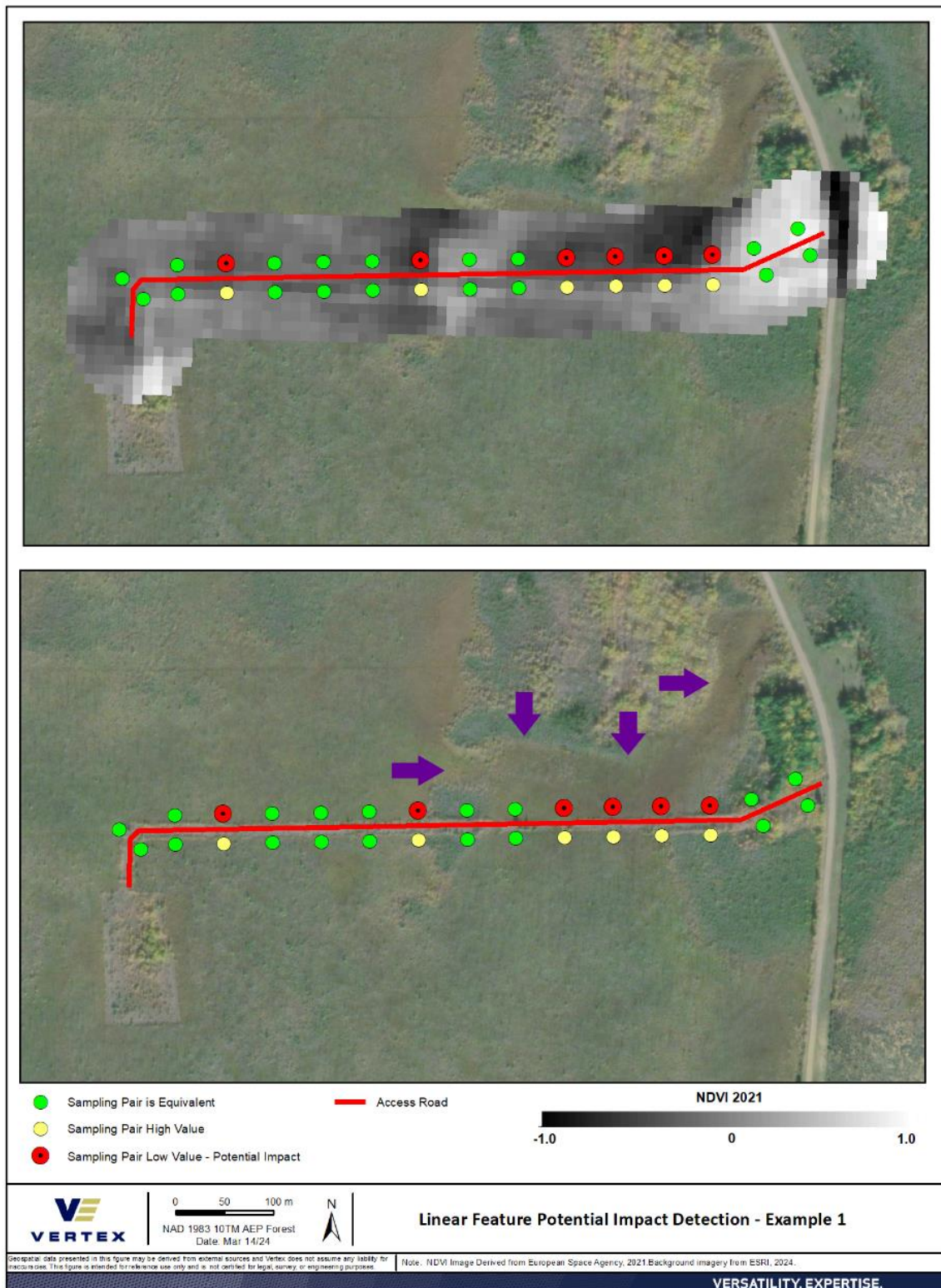
4.3.2 *Outputs*

The following examples show examples of the real-world application of the tools developed here (Figure 11 and Figure 12). We used an interval spacing of 50 m, an offset sampling distance of 15 m, and the difference threshold was set at 15%. In the examples shown, the top frame is the assessed NDVI image with sampling points categorized by NDVI value comparison. The green pairs (one point above and one below the access road indicates a pair, moving along the line) represent areas of minimal to no impact while the yellow/red pairs indicate where one sampled value is more than 15% lower (in red) than the other for value in that pair (in yellow). In the top NDVI image, the lighter grey and white areas in the greyscale image indicate denser or healthier green vegetation. The darker grey to black areas indicate sparse, stressed, or poor vegetation cover, and potentially water. The bottom frame is a reference high spatial resolution image with those same sampling points. The assessed image and reference image may not have been collected in the same year as we relied on the reference imagery supplied by ESRI in their mapping software.

In the first example, four consecutive red sampling points along the top of the line, towards the east, indicate those locations have values lower than the corresponding locations on the southern side of the feature. This indicates a combination of poorer quality vegetation or increased water table effects in this area. The bottom frame shows the same sampling locations where a wetland area is evident as indicated by the purple arrows.

The second example shows potential impacts at several locations along the southern side of the feature. The optical image shows evidence of water ponding in these areas. We also see that several pairs are classified as not having a potential impact, indicated in the bottom frame, even though they are near the visible ponding water. This is probably due to different wetness conditions present in the assessed year (2021) compared to the optical reference year (2020) and to the offset value used (15 m), which was quite low and just at the periphery of the ponding

water, as detected by the satellite. If a larger value had been specified (e.g., 25 m), the undetected values would likely have been classified as potential impacts.



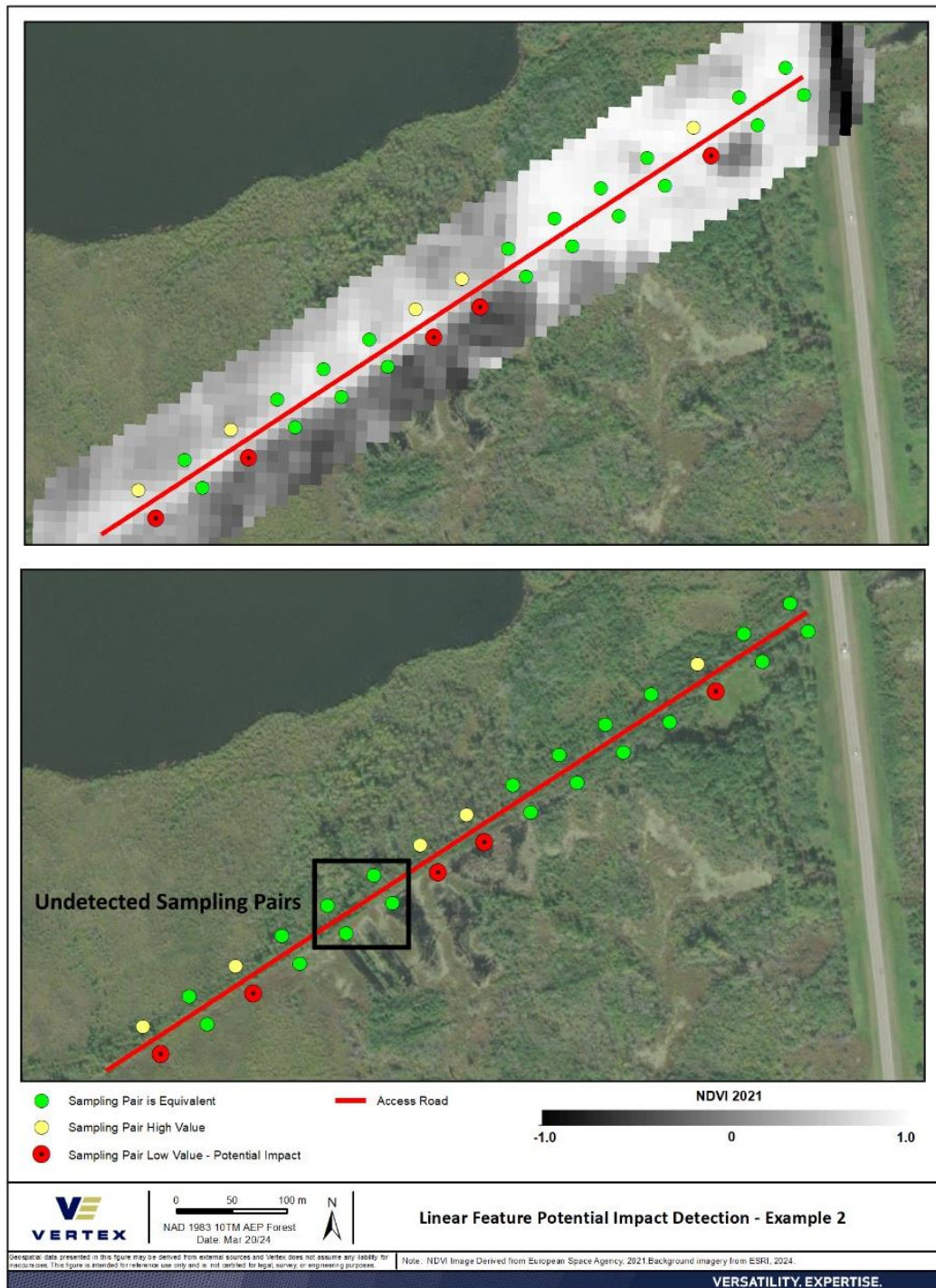


Figure 12. Example 2 from Task 3 output. In this example, we can see that several points along the southern side of the feature are classified as potential impacts. The optical imagery in the bottom frame, from the year prior to the assessed year, shows clear evidence of ponding water. Not all ponds are detected by the sampling pairs, see bottom frame, and there are two probably reasons for this: first is that wetness conditions may have been different between the year of assessment (2021) compared to the optical image (2020), and second, an increase of 1 pixel in the offset distance would have resulted in a positive detection.

4.4 TASK 4

The tools developed for this Task accomplished our goal of implementing a ML-supervised classification algorithm to aid in the differentiation of bogs from fens. This task also facilitated the implementation of Task 1, which was to incorporate *in situ* data with multi-sensor, multi-date EO data. Using training data for tamarack forests and produced from field record-guided image analysis, the land cover classification produced an effective classification of tamarack forest cover.

4.4.1 Tools

A new module was developed in this Task that implements ML-supervised classification of EO imagery. The module includes options for setting classification parameters that guide the classification tool, selection of input layers including multi-temporal optical imagery and SAR imagery, and export of the classified image and related accuracy metrics.

4.4.2 Outputs

The first output of the tool is the map of tamarack forest cover (Figure 13), that shows mapped tamarack forest cover from the supervised classification. The classification identified just over 14 km² of tamarack forest cover, and the associated classification accuracy was 85%. This high accuracy suggests that our methodology creates a reliable map of tamarack forest.

Using this map of tamarack, we then overlaid those forest polygons with bog polygons in the Alberta Merged Wetland Inventory (AMWI) dataset from ABMI (Figure 14). Of the approximately 14 km² tamarack forest, 3.5 km² are within bog areas. This may have potential impact on the decision-making process regarding any padded sites in those areas.

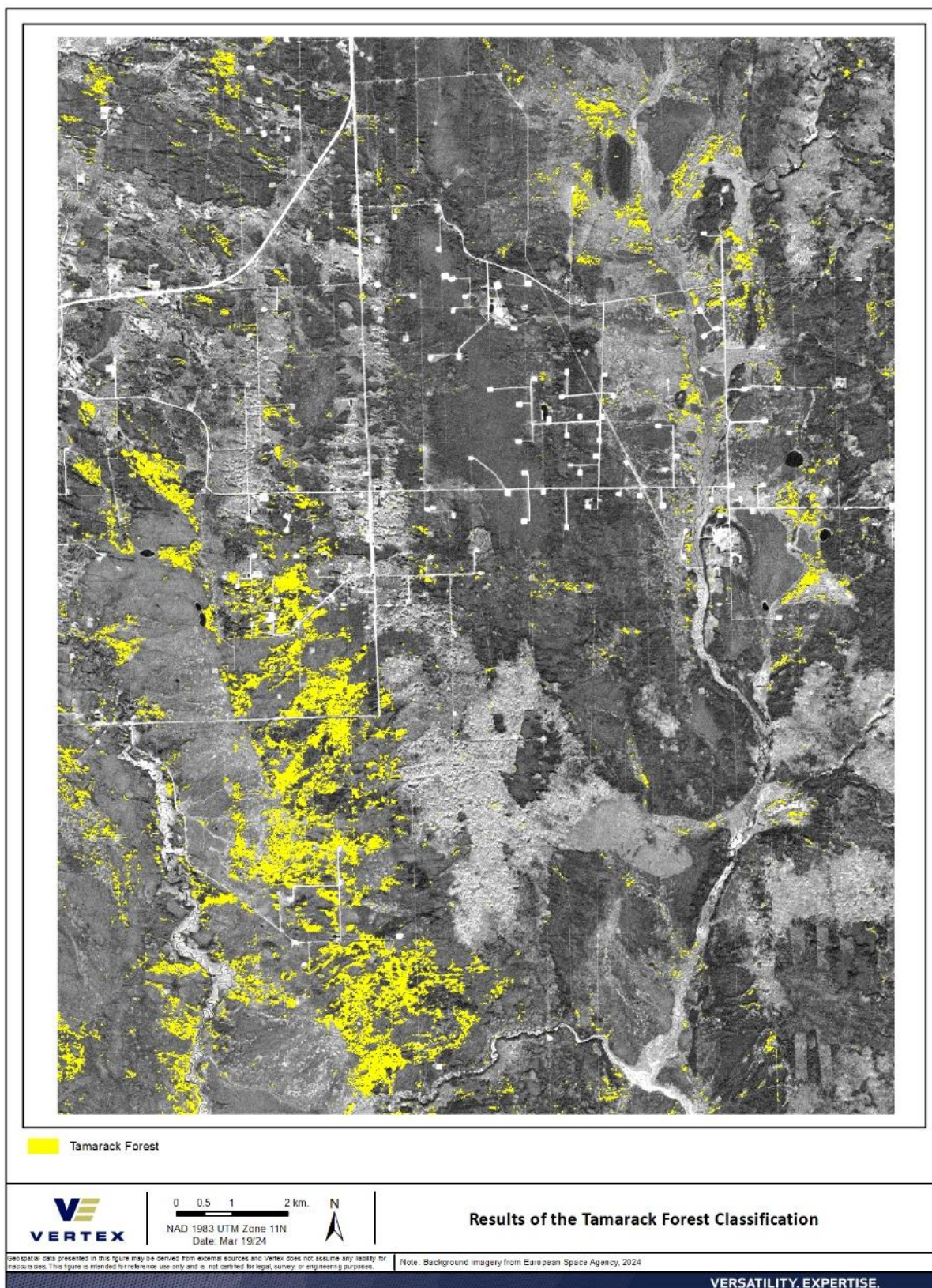


Figure 13. Classification output showing tamarack forest areas detected by the multi-sensor, multi-date supervised classification.

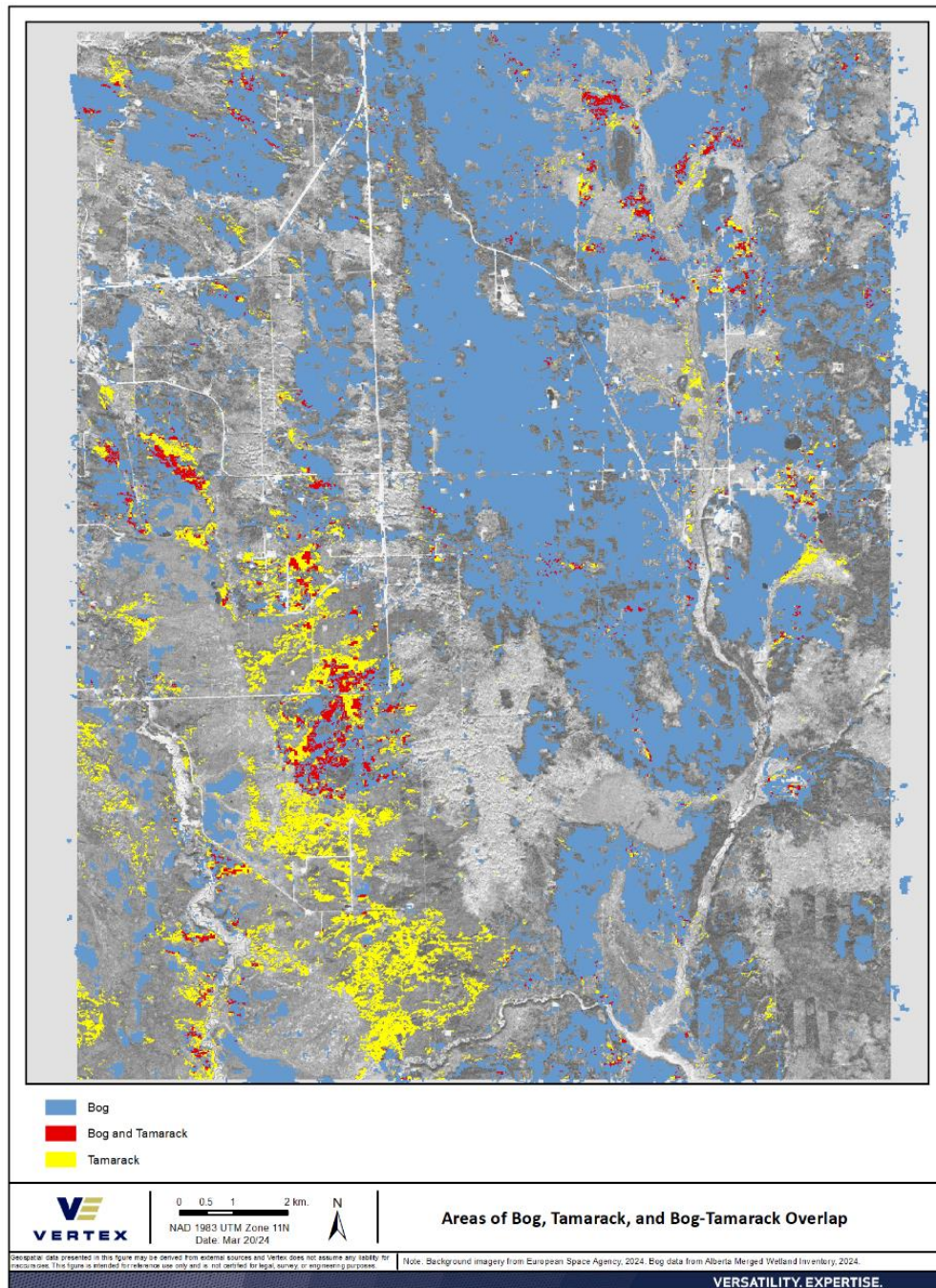


Figure 14. Map of tamarack forest classified using our methods, bogs as mapped by the AMWI, and bog/tamarack overlap areas. We can see that there are several large areas of overlap which could indicate potential misclassification of bogs.

4.5 TASK 5

In this Task, our objectives were to extend our existing tools to unpadded well sites, and to analyse time-series vegetation information to detect disturbances and assess long-term recovery

trends at unpadding sites. We successfully extended our tools to unpadding sites by adapting the padding-site LiDAR processing software to extract forested areas over any input area, padding or unpadding. We were partially successful in our second objective, to analyse time-series data for assessing disturbance and long-term recovery trends.

4.5.1 Tools

- LiDAR processing modules to calculate CHMs from BEDEMs and FFDEMs and to extract forest polygons using a threshold classification method with user-adjustable height parameters.
- Creation of a dense time-series data cube of remote sensing data including the option to perform data smoothing to reduce signal noise.

4.5.2 Outputs

Using the LiDAR tools developed in this Task, we can process large volumes of elevation data and extract forest polygons for a list of input locations. In the example below (Figure 15), we show a sample output from one of the input test locations. The image at left is the CHM derived from the BEDEM subtracted from the FFDEM. On the right, we can see the green polygons that correspond to forested areas where the vegetation canopy is greater than 2 m in height. The forested areas generated here were used in Task 2 to assess forest sustainability.

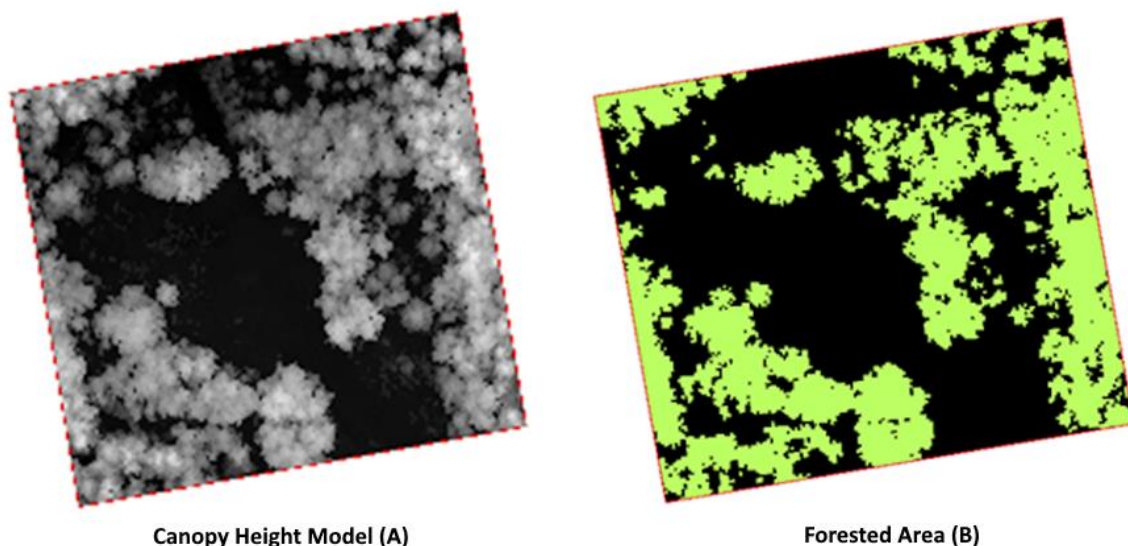


Figure 15. (A) Sample output of the CHM calculated using from LiDAR at native resolution of 0.5m. (B) Resulting forested areas extracted using a minimum height threshold of 2 m.

We successfully developed software to assemble the time-series data cubes and apply a smoothing algorithm to minimise image noise (Figure 16). In the example below, we see the annual NDVI value calculated from a best-pixel analysis, represented by the green line. The black line represents the raw NDVI value trends and the orange line is the NDVI with a smoothing filter applied to reduce data noise. Time constraints prevented us from developing a disturbance detection algorithm to apply to the dense time-series data output.

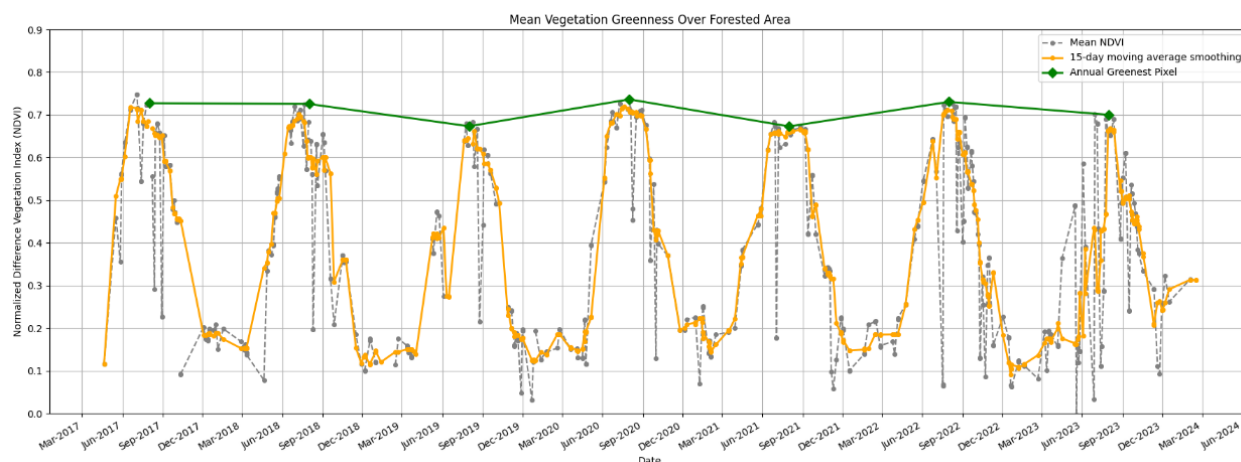


Figure 16. This figure shows the variability of NDVI throughout a dense time series for a 100 m by 100 m test polygon. The green line shows the annual observed NDVI calculated for the entire polygon. The orange line represents the values based on our implementation of a smoothing algorithm designed to remove atmospheric and other noise.

5.0 DISCUSSION

This project was undertaken with the primary objective of developing digital tools for monitoring terrestrial environments impacted by oil and gas operations using EO data. Our efforts resulted in the development several new methods for monitoring forest sustainability, identifying potential hydrologic impacts related to padded linear features, and for mapping indicator species such as tamarack as a potential differentiator of wetland types. We also had success in extending our existing tools to non-padded sites and made progress in using time-series data to identify potential environmental disturbances.

5.1 TASK 1

Using our existing field-based records allowed us to create accurate training polygons for a specific species of interest, a task that is more difficult to accomplish using visual image interpretation only. We created a set of 113 Tamarack Forest polygons guided by the field data which were used in further EO data analysis. This incorporation of field data into a larger remote

sensing analysis program increases the accuracy of the input data for later analysis. While we used field observations to guide training data, various other types of field data, including vegetation species, hydrological information, and other information can potentially be incorporated into the analysis algorithms developed.

5.2 TASK 2

Our most complex objective, from a software development viewpoint, was the implementation of tools to assess the sustainability of forests on developing pads. This Task involved developing a significant number of processing and analysis modules related to LiDAR data manipulation, multi-temporal optical image stacking, cloud-based processing, data extraction and trend analysis. It achieved its goal to establish forest cover polygons in padded sites and extract sustainability trends for those polygons. The analysis provided can be used in several informative ways, including comparing the median NDVI value for the on-pad forest stand to that of undisturbed adjacent forest, which is useful for estimating time-to-recovery. Within the forest polygons we created, the individual pixel-level trend information can highlight any areas of potential concern where vegetation trends are stagnant or declining, even within an overall healthy or positively-trending forest stand. This would allow us to identify potential barriers to sustainability or areas where further remediation may be desired.

Significant time savings can be realized by using this method of forest sustainability assessment. Our testing dataset of 20 sites required less than a day for initial data preparation, software parameterization, processing time, and reporting. Extending the time required, we anticipate the ability to analyse hundreds of sites in only a matter of days, compared to potentially months or more that would be required by standard field assessments. Those traditional field assessments would also be challenged to provide historical vegetation change record, limiting their utility for assessing long-term sustainability.

The tools we have developed are flexible in terms of the time-series data construction and for the EO data types included in the analysis, allowing it to adapt to longer time-series datasets or to assess other types of sites beyond padded sites. In addition, the LiDAR processing tools can also be used to detect features other than forest including low vegetation cover or exposed areas.

The tool works as intended though we have identified several areas for further development. In many instances, we observed oddly shaped polygons, such as we see in the red polygon in Figure 9. This is an artifact caused by the footprint of the site being misaligned with the LiDAR imagery and may result in undisturbed forests being delineated. Another issue we encountered relates to the method used to extract image pixels based on irregular polygons. We tried several methods and each one presented different challenges regarding inclusion or exclusion of forested pixels. Finally, in Figure 10, we can see several small polygons that may not represent relatively large forest stands therefore might bias the site results.

5.3 TASK 3

Padded sites and related padded linear features can impact peatlands by altering natural hydrological process in a negative way. In Task 3 we developed methods to assess the impact of padded features on peatlands. We accomplished this by developing a set of tools that allow us to assess a large volume of linear features for potential impacts. We created a set of access roads, automated the creation of sampling locations along those roads, and then identified potential impacts in a much faster and more widely distributed way than would be otherwise achievable in a traditional field campaign. The analysis tools developed detected changes in vegetation and wetness conditions along the input features and identified areas of potential impact for further investigation. This investigation includes an initial visual inspection to determine whether the impact is due to natural landscape variability or to human disturbance. In areas of genuine impact, the outputs created from our methods can be used to set up a longer-term remote sensing monitoring program, to plan future field work to verify results and look for root causes of impact, or as a planning tool to organize additional remediation efforts and allocate resources efficiently. As with the forest sustainability tools, they can be easily extended beyond the initial application. For example, we could use the tools to assess pipelines and compare the values of vegetation metrics on the disturbed pipeline to undisturbed vegetation along the feature. The sampling intervals and offsets are data- and need-driven, so we can sample as frequently or sparsely as is needed to accomplish the task using an array of spectral indices that can address wetness, soil moisture, leaf area, and other vegetation parameters.

Some of the efficiency realised using EO data is the ability to quickly assess large areas compared to the time needed for a traditional field campaign. To illustrate this, our test dataset consisted of 81 km of access roads which were distributed over an area of nearly 20,000 km². We processed these features to generate 3,120 sampling points. Using the input linear feature dataset and processing of historical data, we sampled this entire area and created the classified assessment points in less than a day. As with the forest sustainability tool created in Task 2, this is a significant improvement compared to a traditional field sampling program to achieve the same results. Using these results of impacted areas greatly enhances the ability to plan validation fieldwork, ground-based investigations of the root causes, or to implement mitigation strategies. The results can also be segregated into smaller datasets that can be continually monitored using the same open source imagery, or commercial imagery can leveraged to provide greater information without the cost of acquiring large volumes of imagery.

5.4 TASK 4

We used the *in situ* derived data from Task 1 to guide a supervised classification of the potential tamarack forests. This was implemented with a newly developed tool that builds single-date or

multi-date optical image data cubes, along with any desired spectral indices, and incorporates SAR data as required. These data were classified using a ML algorithm into various land cover classes that were then analysed for their ability to differentiate tamarack forests in relation to wetland bogs in attempt to enhance the accuracy of fen and bog differentiation. The tools and training data from Task 1 resulted in a high-accuracy classification (85%) of tamarack forest. Overlay with existing bog maps showed several areas of high tamarack forest cover. Further investigation into the regions identified as having significant overlap would be required to field-validate the results of the classification.

5.5 Task 5

The underlying idea behind the adaptation of existing tools to unpadded sites was to be efficient in the development of new capabilities. Using the LiDAR tools already completed allowed us to create the forested area mapping tools and expand the application of that original toolset to padded or unpadded sites.

The dense time-series data analysis tools could provide valuable insight into short term impacts that are missed with an annual or single-date analysis of a study site or sites, though testing was not able to be conducted at this time.

6.0 CONCLUSIONS

The project met its desired outcomes to further the application of remote sensing technologies to monitor the impacts of padded and unpadded sites in wetland areas. We added significant capacity to assess several types of impact and identify areas of concern, such as:

- We were able to update existing tools and develop a significant number of new tools for monitoring forest sustainability, impacts to wetlands from padded features, and to steps towards improving bog and fen differentiation.
- While the specific algorithms described here relate to four specific tasks, most of the code employed is modular in nature and allow these tools to be adapted for similar uses or entirely new ones.
- These tools allow us to sample vast areas in very short amount of time.
- We can also sample in areas where access is prohibited due terrain or land cover challenges, private land ownership, or other impediments.
- We can now analyse hundreds of sites for different types of impact including vegetation health, moisture or wetness, and impacts to surrounding wetlands. We can do this over tightly grouped sites in one localized area or for sites distributed over the entire province. Depending on the number of sites and metrics required, we can complete the analysis in as little as a few days, greatly enhancing the ability to plan future validation, monitoring, or remediation work.
- We implemented our code base in such a way that we can readily expand the tools beyond their primary interests to further help understand impacts or reclamation status of oil and gas related development. This greatly increases their value.

7.0 RECOMMENDATIONS

Further work is recommended in several areas to build upon the capacity developed in this project by improving various algorithms, validating results using future field work, and enhancing the reporting and communication methods of the resulting analysis. Below are recommendations as they relate to the objectives of the individual Tasks:

7.1 TASK 1

- Explore options for incorporating validation of the tamarack results with ongoing field work.
- Continue to evaluate options for incorporating existing and planned field observations with remote sensing analysis.

7.2 TASK 2

- Continue development of reporting and communication methods to improve readability and reduce the time required for reporting.
- Expand testing to higher resolution optical data that can take advantage of the higher resolution available in the LiDAR imagery. PlanetLabs or SPOT satellite imagery could be useful for the spectral or temporal resolutions, respectively.
- Use alternative metrics for comparison of trends and target vegetation goals to indicate sustainability or recovery success.
- We experimented with several methods to extract image pixels based on irregular polygons and each method presented challenges to inclusion or exclusion of pixels – we could explore alternative methods.

7.3 TASK 3

- Continue work on reporting and communication methods to improve readability.
- Using established impacts, we can work on additional tools to work backwards in time to potentially derive the time impacts began, and thus narrowing the potential area requiring remedial action.
- Expand the sampling offsets in areas of interest to include the ability to detect the magnitude, or distance away from the linear feature, of the impact.
- Expand the sampling program to assess on-feature conditions compared to along-feature conditions. This could be useful for estimating vegetation conditions along reclaimed linear features such as pipelines.

7.4 TASK 4

- Conduct field assessments of the areas of significant tamarack and bog overlap to assess the nature of the wetland area.
- Expand the classification to an additional area where padded sites are found in bogs, based on our previous work in padded site identification.

7.5 TASK 5

- Continue development of the dense time-series tool by selecting unpadded sites that have potential water ponding impacts.
- Develop detection algorithms to automatically assess sites where water ponding can be detected.

8.0 APPLICATION

Our target market are producers and service companies within the oil and gas industry requiring environmental monitoring services for any stage of well site development. Our end users will be predominately oil and gas firms, the Orphan Well Association, pipeline companies, and consulting and engineering firms. Vertex will leverage its large network of industry partners to engage with oil and gas companies. Vertex will connect with other users through social media, networking events, and by presenting at industry or governmental conferences. We further anticipate other consulting companies will rely on external remote sensing services because very few companies have dedicated remote sensing staff capable of conducting EO data analysis in a cost-effective manner.

Vertex has worked in the environmental monitoring sector for oil and gas industry for over 30 years and has intimate knowledge of the market demand. We work with small, medium, and large producers and all have reiterated that they are looking for ways to conduct environmental monitoring faster, better, and cheaper. Industry is increasingly interested in remote sensing technologies and the efficiencies they enable on the path to final site reclamation certification. We will offer these newly developed tools both as a fee-for-service in our existing portfolio, as well as a subscription-based service for longer-term monitoring.

9.0 REFERENCES

- Abacus Datagraphics Ltd. 2022. AbaData 2.0 [Software]. <https://abadata.ca/>
- Alberta Agriculture and Forestry. 2020. Derived Ecosite Phase version 2.0 [Data file]. Alberta Agriculture and Forestry, Edmonton, Alberta. <https://open.alberta.ca/publications/derived-ecosite-phase-version-2>
- Alberta Biodiversity Monitoring Institute. 2021. ABMI Wetland Inventory. [Data file]. <https://abmi.ca/home/data-analytics/da-top/da-product-overview/Advanced-Landcover-Prediction-and-Habitat-Assessment--ALPHA--Products/ABMI-Wetland-Inventory.html>
- Alberta Energy Regulator. 2022. Digital Integrated Dispositions. [Data file]. <https://www.alberta.ca/digital-integrated-dispositions.aspx>
- Caron, V., M. Henley, E. Loos, K. Renkema, D. MacKenzie and N. Shelby-James. 2022. Identification and Characterization of Abandoned Padded Well Sites Using Remote Sensing. Prepared for the Petroleum Technology Alliance of Canada, Calgary, Alberta. Report 20-RRRC-05_1. 17 pp.
- European Space Agency. 2022. Copernicus Open Access Hub. [Web portal]. <https://scihub.copernicus.eu/>
- Government of Alberta. 2015. Alberta Wetland Classification System. Accessed online <https://open.alberta.ca/dataset/92fbfbf5-62e1-49c7-aa13-8970a099f97d/resource/1e4372ca-b99c-4990-b4f5-dbac23424e3a/download/2015-Alberta-Wetland-Classification-System-June-01-2015.pdf>.