Technical Memo: Riparian Planting Survival Assessment

Prepared for

Ministry for the Environment

March 2023

PATTLE DELAMORE PARTNERS LTD
Level 4, 111 Customhouse Quay
Wellington 6011
PO Box 6136, Wellington 6141, New Zealand

Tel +64 4 **471 4130**
Website **http://www.pdp.co.nz**
Auckland **Hamilton** Tauranga **Wellington**
Christchurch Invercargill

solutions for your environment

Quality Control Sheet

TITLE	Technical Memo: Riparian Planting Survival Assessment
CLIENT	Ministry for the Environment
VERSION	Final
ISSUE DATE	7 March 2023
JOB REFERENCE	W02422800
SOURCE FILE(S)	W02422800_Stage2_TechnicalMemo-draft

DOCUMENT CONTRIBUTORS

Prepared by

SIGNATURE

Marinus Boon

David Knox (Lynker Analytics)

Reviewed by

SIGNATURE

Matt Lythe (Lynker Analytics)

Approved by

David Whitty

Limitations:

This report has been prepared by Pattle Delamore Partners Limited (PDP) on the basis of information provided by the Ministry for the Environment [and] [others (not directly contracted by PDP for the work)], including Land Information New Zealand and Maxar SECUREWATCH data supplied by MfE. PDP has not independently verified the provided information and has relied upon it being accurate and sufficient for use by PDP in preparing the report. PDP accepts no responsibility for errors or omissions in, or the currency or sufficiency of, the provided information.

This report has been prepared by PDP on the specific instructions of the Ministry for the Environment for the limited purposes described in the report. PDP accepts no liability if the report is used for a different purpose or if it is used or relied on by any other person. Any such use or reliance will be solely at their own risk.

Executive Summary

Riparian areas are the strips of land beside drains, streams, rivers, and lakes. They include areas on-farm where the soils are wettest, such as wetlands, springs or seeps and gullies. Riparian planting has multiple environmental benefits including water filtration, erosion prevention, moderation of water flow, shading waterways, providing habitat for indigenous species, and keeping livestock out of waterways.

Recently riparian planting has received a major boost through the Ministry's Mahi mō te Taiao | Jobs for Nature funding programme. This programme manages funding across multiple government agencies to benefit the environment, people and the regions and will run until 2025, as part of the COVID-19 recovery package.

Under this contract, the PDP/Lynker team assessed the suitability of a range of remote sensing methods to monitor riparian ecosystems for 5 pilot sites. This included satellite, airborne and unmanned aerial systems. For each sensor, the team developed machine learning techniques to identify and map the extent of plants within riparian systems. Stage 1 concluded that high resolution Maxar satellite imagery is the most suitable.

Stage 2 involved the implementation of a Maxar model over 141 sites in the Taranaki region. The final prediction model reported an overall classification accuracy of 72% against the validation data, however, with masking known waterways the model accuracy rises to 80% with unvegetated, woody and grass classes detected with F1 scores between 0.8 and 0.9. The model had difficulty resolving narrow features such as streams and herbaceous strips which affected the unvegetated and herbaceous classes. Masking of waterways, including narrow streams boosts the accuracy considerably and is recommended as a pre-step. However, the overall land cover representation is good.

In conclusion, the developed model achieves the project objectives of determining riparian planting survival with an acceptable accuracy, mapping riparian vegetation into 6 classes. The method can be applied to predict riparian planting survival across New Zealand.

Table of Contents

Table of Figures

Table of Tables

Appendices

1.0 Purpose and Scope

1.1 Purpose

The Ministry for the Environment (MfE, the Ministry) engaged Pattle Delamore Partners (PDP) and Lynker Analytics (Lynker) to develop a computer-based model using machine-learning to establish a proof of concept for the Jobs for Nature environmental monitoring programme of work run by the Ministry. The Ministry is seeking to have an independent mechanism to assess the survival of funded riparian planting for assurance and impact evaluation purposes. The project has tested methods to meet those requirements.

The Ministry wants to test the suitability of different types and resolutions of imagery to determine the survival of various plant species at different life stages with acceptable accuracy and precision, and has requested advice on how to upscale this analysis from pilot sites to regional and national programmes of work.

The project encompassed:

- Imagery data acquisition and collation for pilot sites;
- Classification and analysis of riparian plants identified at multiple sites using multiple types of imagery;
- Field validation of plant classification outputs;
- Advice on upscaling the plant survival “chain of analysis” to a regional and national level.

This document has been prepared by Pattle Delamore Partners, Lynker Analytics and with input from Papawera Geological Consultants (Karen Denyer).

1.2 Scope

To facilitate the delivery of the above, the project was split into 3 distinct stages:

- Stage 1; pilot sites analysis
- Stage 2; all sites analysis
- Stage 3; future focused evaluation

This report addresses the scope and deliverables of Stage 1 and 2, which is broken into further components.

2.0 Stage 1 Pilot Site Analysis

This stage included the development of riparian planting survival plant classification models using machine learning for imagery analysis. The delivery of these models included several steps including the development of models, imagery acquisition, application of the models, field validation and data delivery. Model development involved the establishment of training classes, an iterative process of model refinement including field validation of riparian vegetation classes.

Models have been developed for four different sensors:

- Medium-resolution satellite imagery;
- Aerial imagery;
- High-resolution satellite imagery; and
- Unmanned Aerial Vehicle (UAV).

2.1 Stage 1 Methods

The Ministry has requested the following categories of information to be incorporated into the assessment of riparian planting success (assessed within an 18 to 60-month post-planting timeframe):

1. Woody
2. Herbaceous
3. Sedges
4. Grasses
5. Vegetation height (tall, medium, short, i.e., 5m+, 1-5 m, <1 m)
6. Canopy closure (sporadic, closed - proportion undefined by MfE).

In addition, the Ministry requested a differentiation between ‘intended’ and ‘unintended’ plants, with ‘intended’ plants being defined as those funded for planting versus those unlikely to be funded for planting (‘unintended’ plants). The Ministry requested that an “intended to unintended ratio” be the basis for determining plant survival rates.

Because some unintended plants (i.e., self-seeded) may be ecologically appropriate and indistinguishable from planted plants, it is proposed to use the term “desirable” as an aggregate category of indigenous species that are naturally occurring in the area and typically grow in riparian zones.

The first four of the MfE categories were compressed into a single ‘Vegetation Structure’ category (attribute) with woody, herbaceous, sedgeland and grassland as potential classes (note that these vegetation structures do not imply indigenous vegetation). Within any given planned planting zone, unvegetated land covers may also be legitimately present (e.g., water, hard surfaces such as walkways, and bridges), and these were incorporated into the Vegetation Structure attribute under the 'unvegetated' class.

Broadly, measures of riparian planting success could look like Table 1 after 5 years. All planted areas are expected to eventually reach 80-100% canopy closure, to shade out and exclude undesirable species (weeds).

Table 1: Potential indicators of successful planting after 5 years

	Successful planting value		
Vegetation Structure (MfE categories 1-4	**Height (MfE category 5)**	**Canopy closure (MfE category 6)**	**Coverage of desirable[1] species**
Woody	Medium-tall	Closed	> target %
Herbaceous	Short-medium	Closed	> target %
Sedgeland	Short-medium	Closed	> target %
Grassland	n/a grassland likely represents failed planting as true native sward-forming grasses are unlikely to be planted in NZ riparian projects, however narrow strips of rank grass are often an intended outcome to protect fences and trap silt	n/a grassland likely represents failed planting as true native sward-forming grasses are unlikely to be planted in NZ riparian projects, however narrow strips of rank grass are often an intended outcome to protect fences and trap silt	n/a

Notes:

1. *Desirable species are either those deliberately planted (intended species) or self-seeded native plants that are ecologically appropriate to the local area and riparian zone.*

These 5 categories are not mutually exclusive, however, and younger plantings cannot be expected to reach the levels of canopy height or closure depicted in Table 2. Because suitable imagery for assessment may not match the ideal 5-year time frame to allow riparian planting to reach its full potential, a more nuanced approach is proposed as presented in Table 2.

Expected Canopy Closure values are proposed for three age classes, to allow for an acceptability measure based on time since planting. These target values are based on best practice plant spacing (0.5-1 m for herbaceous species, 1-1.5 m for shrub/small tree species). Wider spacing may lead to a failed riparian planting by taking longer than 5 years to reach canopy closure and facilitating unintended plant establishment.

Table 2: Proposed indicators of successful planting at different age stages based on best practice spacing

Actual vegetation Structure	Months since planting	Expected Height	Expected Canopy Closure	Coverage of desirable[1] species	Comments
Woody	5+ years (>60 months)	Medium or tall	>80%	>72%	If planted at <=1 m spacing a woody canopy should be closed after 5 years. As plants are generally circular there will be natural gaps - this is allowed for by setting 80% closure as success measure for plants > 5 m tall.
Woody	2-5 years (24-60 months	Medium	50-80%	45-72%	At less than 5 years old, trees and shrubs will likely be less than 5m tall. Some shrub species are slow to fill out, e.g., mānuka, small leaved coprosmas and tend to grow up more than out.
Woody	< 2 years (<24 months)	Short	10-50%	9-45%	At < 1 m and planted at 0.5 to 1 m spacing, shrubs like mānuka, tī kōuka, hebe, and karamū are probably less than 30cm in foliage diameter and may cover little more than 10% of a planting zone.
Herbaceous/ Sedgeland	5+ years (>60 months)	Short to medium	>80%	>80%	At 1 m spacing, canopy should be >80% if plants are taller than 1 m. Narrower plants like *Macherina* rushes will take longer to reach canopy closure
Herbaceous/ Sedgeland	2-5 years (24-60 months	Short to medium	50-80%	50-80%	At planting, *Carex secta/Cyperus* etc are likely 50 cm tall, 20-30 cm across. They are usually planted after 1 year likely 60-70 cm across, after 2 years 1 m across.
Herbaceous/ Sedgeland	< 2 years (<24 months)	Short	10-50%	10-50%	Assuming these plants are less than 2 years old, or short stature plants (e.g., *Carex maorica*), or were much smaller when planted. Such

Table 2: Proposed indicators of successful planting at different age stages based on best practice spacing

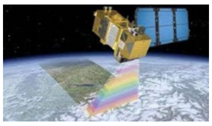
Actual vegetation Structure	Months since planting	Expected Height	Expected Canopy Closure	Coverage of desirable[1] species	Comments
					short plants may be only 30 cm across after a year.
Grassland	0-5+ years	Short	n/a	n/a	
Unvegetated	0-5+ years	Short	n/a	n/a	

Notes:

1. *These will need field checking and refining to ensure they are reasonable values. They will also depend on size and spacing of plants at the time of planting.*
2. *Desirable species are either those deliberately planted (intended species) or self-seeded native plants that are ecologically appropriate to the local area and riparian zone. This proposal allows for undesirable/unintended species to comprise up to 10% of that portion of the planned planting zone no longer in residual grassland/bare ground. Thus if 50% of the planned planting zone (PPZ) is no longer in grassland or bare ground, it is acceptable for up to 5% of the total PPZ to be in unintended vegetation. If 80% of the PPZ is no longer in grassland or bare ground, it is acceptable for 8% of the PPZ to be in unintended vegetation.*

Imagery acquisition and plant classification models were developed for 5 pilot sites as indicated in Figure 1. Selection of these sites involved looking at various criteria including planting age (sites >=60 months or nearest to 60 months) especially in relation to the latest available regional aerial imagery (age at image date), different regional characteristics, logistics for UAV imagery collection, access, and planting information. Table 3 provides an overview of all the data inputs used for the Stage 1 pilot site analysis. The Stage 1 analysis included the use of only true colour RGB (Red Green Blue) imagery. Table 4 provides an overview of the pilot sites location, planting year, imagery acquisition dates, planting and ecosystem types.

Table 3: Overview of data inputs

<table>
<tr><th>Data set</th><th>Description</th></tr>
<tr><td>Medium-resolution satellite imagery
</td><td>Sentinel 2 Copernicus RGB composite (acquisition dates ranging from 02 2019 for Paramata to 04 and 05 2021 for other sites) at 10m resolution.

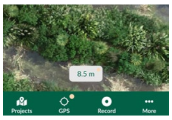
Aerial imagery training was used to train the Sentinel 2 model.</td></tr>
<tr><td>Aerial imagery
</td><td>Regional and MfE LUCAS Deforestation Mapping project (for Taranaki site) RGB imagery down sampled to 30 cm ranging from 2016-2021 summer imagery.</td></tr>
<tr><td>High-resolution satellite imagery
</td><td>Maxar pansharpened RGB imagery (panchromatic band at 0.5m resolution and multispectral at 2.0m) acquired for summer dates between 2020-2.</td></tr>
<tr><td>Unmanned Aerial Vehicle (UAV)
</td><td>High-resolution georeferenced RGB imagery down sampled to 0.05 m. Imagery acquired during summer 2021/22</td></tr>
<tr><td>Training data
</td><td>Human annotated training data generated from Aerial, Maxar satellite and UAV imagery using the Active Learning Annotation System (ALAS).

<table>
<tr><th colspan="2">Plant classification used per model</th></tr>
<tr><th>S2, Aerial and Maxar</th><th>UAV</th></tr>
<tr><td>Woody</td><td>Woody</td></tr>
<tr><td>Herbaceous</td><td>Herbaceous</td></tr>
<tr><td>Grasses</td><td>Sedges</td></tr>
<tr><td>Unvegetated</td><td>Grasses</td></tr>
<tr><td></td><td>Unvegetated</td></tr>
<tr><td></td><td>Undesirable herbaceous</td></tr>
</table>

⅓ randomly selected to measure model performance

⅔ sites directly used by the machine learning model</td></tr>
</table>

Table 3: Overview of data inputs

Data set	Description
Validation data 	Field validation data captured using a mobile field data collection application to test the model performance (not used to train the machine-learning model).
Planting data	MfE and Horizons Regional Council (HRC) planting data.

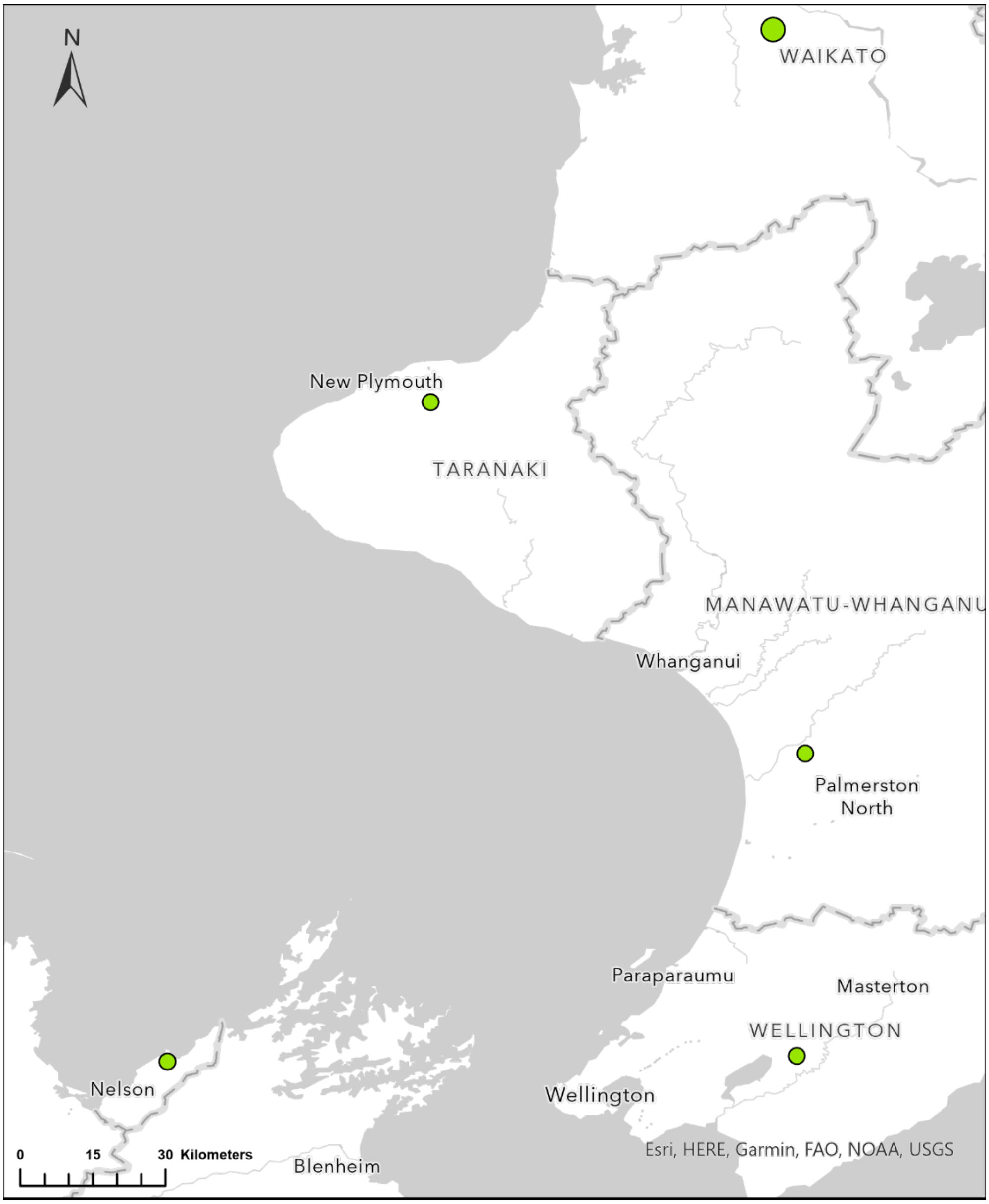


Figure 1 Location of the Stage 1 pilot sites.

Table 4: Overview of pilot sites shown in Figure 1

Pilot site	Region	Planting year	Imagery date	Planting type	Ecosystem
Rotopiko	Waikato	2012-2021 (ongoing)	S2[1]:20210503 A[2]: 2016 and 2019 M[3]:20210324 U[4]:2022019	Woody, Sedges and herbaceous	Restored Peat Lake
Taranaki	Taranaki	2018 and 2019	S2: 20210428 A:2020 M:20210131 U:20220121	Woody	Fenced Riparian area along stream on working Dairy farm
Sanson	Horizons (Manawatu-Wanganui)	2015	S2: 20210525 A: 20210831 M:20210831 U:20220326	Woody and herbaceous	Fenced Riparian area along stream on working Dairy farm
Kaiwaiwai	Greater Wellington	2016 and 2018	S2: 20210525 A:2017 and 2021 M:20210324 U:20220429	Woody, sedges and herbaceous	Constructed wetland on working Dairy farm
Paremata Flats	Nelson	2015	S2: 20190203 A: 2018 and 2019 M:20200202 U:20220316	Woody	Tidal river and land riparian area

Notes:

1. *S2: Sentinel 2 satellite*
2. *A: Regional aerial imagery*
3. *M: Maxar satellite imagery*
4. *U: UAV imagery*

We established a model training environment (Figure 2) called the Active Learning Annotation System (ALAS) with plant classification for all four of the sensors. This system is an image classifier where the machine learning model learns in real-time by presenting images for the trainer to classify. This step included an iterative process of testing, reviewing, and deciding on the classes for each model (sensor). A four-class classification was selected for the medium-resolution satellite imagery; aerial imagery; high-resolution satellite and a 6 class for the UAV imagery for the final Stage 1 models.

Stage 1 Machine Learning Methods

Two deep convolutional neural network architectures for image segmentation were tried across the different sensor types.

Active learning allows us to build up a training library of images and annotations iteratively with model training, thus we target the hardest to classify examples in our training and build up an efficient training set of data. This iterative model training method is shown in Figure 2. The training images and annotations can then be used to train additional models. In Phase 1, we trained models of two types; a patch segmentation model, based on the InceptionV3 CNN model as seen in Figure 4 and a semantic segmentation model using DeepLabV3+.

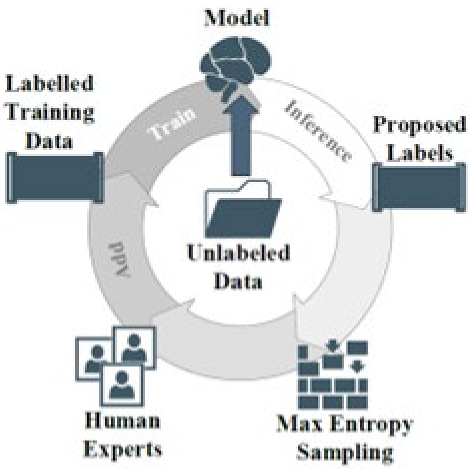


Figure 2: Active Learning Flowchart

In all cases, the machine learning models were trained on imagery and annotations taken from separate training sites, and the test site data was not seen by the models during training. All machine learning models for Stage 1 used the RGB channels only as the deep convolutional neural network (CNN) architectures used were designed for RGB imagery. This also allowed us to use pretrained models that were previously trained on a large image classification task, imagenet, prior to our task. Using pretrained models allows us to finetune our models which requires less data to achieve higher accuracy.

The Active Learning Annotation System (Figure 3) trains a patch segmentation system shown in Figure 4. Models labelled as "ALAS" are patch segmentation models using the CNN architecture, trained using the active learning technique.

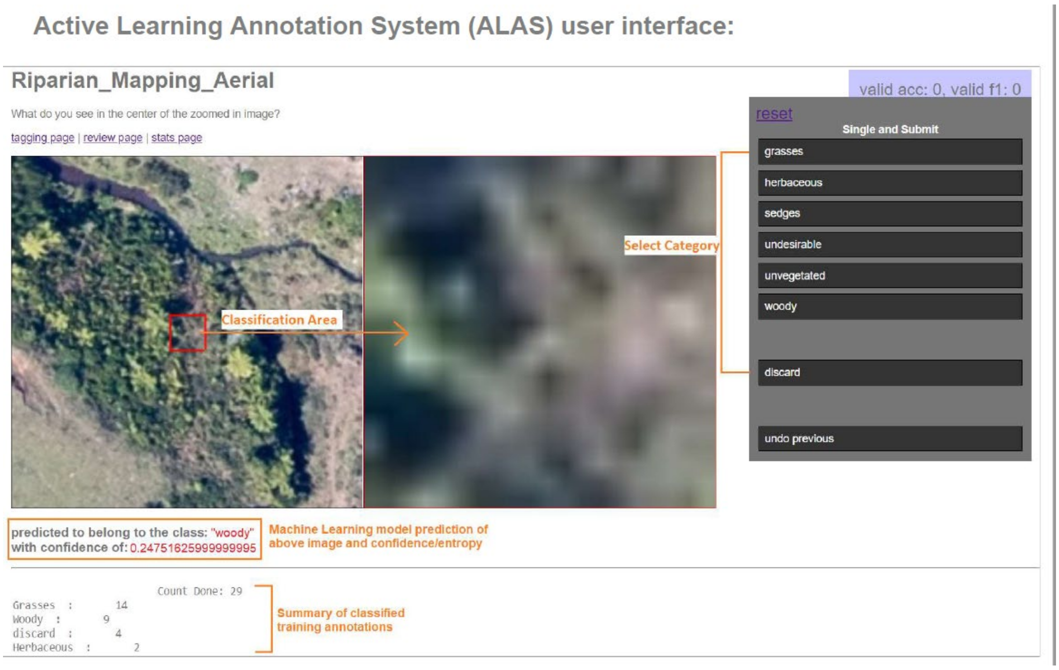


Figure 3: Screenshot of the Active Learning Annotation System (ALAS)

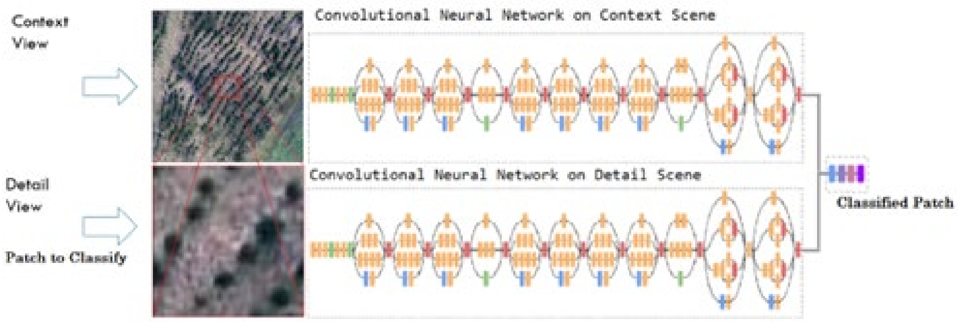


Figure 4: Patch Segmentation CNN architecture

<u>Field Validation</u>

A plant classification schema for the field validation component was developed and implemented using the Input application (Mergin Maps) developed by Lutra Consulting Limited. Riparian planting data was collected in the field using Input on an Android or iPhone/iPad device for all five of the pilot sites. The field schema included detailed vegetation categories to adequately record riparian survival in the field, although provision was made to merge classes for validation against the MfE categories. Figure 5 presents screenshots of the field app indicating the features collected during field validation.

Input (Mergin Maps) was selected as it is free, open-source software (https://play.google.com/store/apps/details?id=uk.co.lutraconsulting&hl=en&gl=US&pli=1 & https://merginmaps.com/start-for-free). High-resolution imagery layers can be uploaded for use in the field, enabling offline use, and additional users can be added without the need for any further user/organisational licences. It has a built-in service for synchronisation and storing data, and users can easily create, transfer and manage their survey project for field data collection.

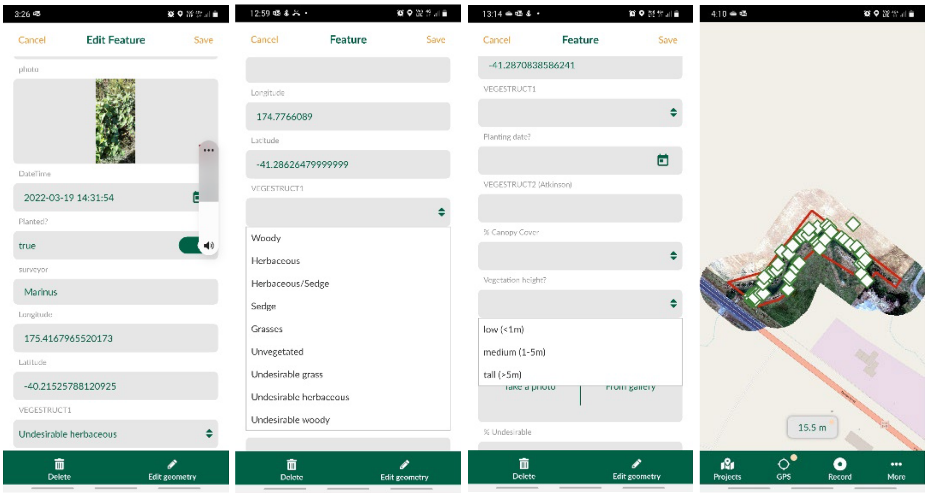


Figure 5: Screenshots of the field validation application

Stage 1 data delivery

Data delivery was completed in accordance with MfE requirements and was delivered as an Esri File geodatabase in NZTM 2000 projection.

2.2 Stage 1 Results

Table 5 presents the mean accuracy of the machine learning model relative to ground checked validation data points across all the classes. Maxar had the highest average mean accuracy score across all the sensors, followed by UAV, based on model results for all 5 pilot sites. However, it is notable that the mean Maxar accuracies for Rotopiko and Taranaki sites is much lower compared to the 3 other pilot sites. These two sites had large heterogeneity in terms of distribution of patches woody and herbaceous cover. Correct prediction of smaller patches of vegetation was challenging for the model since these are below the resolution of the model.

Table 5: Mean accuracy relative to field validation

Site	S2	Maxar	Aerial	UAV
Rotopiko	0.60	0.55	0.66	0.69
Taranaki	0.76	0.57	0.77	0.53
Sanson	0.13	0.92	0.23	0.50
Kaiwaiwai	0.74	0.89	0.78	0.83
Paremata	0.91	0.85	0.92	0.89
Average	0.63	0.76	0.67	0.69

Table 6 presents the mean accuracy and F1 scores across all classes relative to the desktop validation results. The values presented are the mean value across all validation sites. An F1 score is the weighted average of the machine learning precision and recall metrics, and reaches its best value at 1 and worst value at 0. These accuracy results suggest that the UAV model had the highest average accuracy results followed by Maxar.

We chose the F1 score as a measure of accuracy per class as this method strikes a good balance between recall and precision (being the harmonic mean of precision and recall) and is a good classification accuracy measure when data may have more samples of some classes than others. In our case, efforts were made in the desktop validation to select an even number of samples in each class.

Table 6: Mean F1 across all the classes

Sensor	Herbaceous F1	Unvegetated F1	Woody F1	Grasses F1	Average
S2	0.23	0.08	0.27	0.41	0.25
Maxar	0.58	0.25	0.61	0.63	0.51
Aerial	0.31	0.41	0.67	0.55	0.49
UAV	0.30	0.76	0.77	0.79	0.66

Notes:

1. *F1 is the weighted average precision and recall scores - each weighted by the number of points with that class.vs Desktop*

A summary of the 'intended' to 'unintended' plant ratio of the five pilot sites for all sensors is presented in Table 7. 'Unintended' is defined as 'undesirable' and 'grass' classes only. 'For 'unintended' calculations, 'undesirable' was only used for the UAV datasets due to resolution limitations, and the remaining datasets used 'grass'. 'Total Cover' is defined all classes except 'unvegetated'.

Table 7: Intended and unintended cover

Sensor	Site	Unintended (ha)	Unintended % of Total Cover	Intended (ha)	Intended % of Total Cover	Total Cover (ha)
S2	Horizons	0.53	92.35	0.04	7.65	0.57
	Kaiwaiwai	0.23	21.31	0.85	78.69	1.08
	Paremata	0.04	2.76	1.32	97.24	1.36
	Rotopiko	2.46	16.44	12.52	83.56	14.98
	Taranaki	0.36	30.74	0.80	69.26	1.16
Maxar	Horizons	0.03	11.03	0.24	88.97	0.27
	Kaiwaiwai	0.03	3.63	0.75	96.37	0.78
	Paremata	0.09	7.85	1.00	92.15	1.09
	Rotopiko	3.56	23.41	11.64	76.59	15.19
	Taranaki	0.34	34.12	0.66	65.88	1.00
Aerial	Horizons	0.47	86.31	0.07	13.69	0.54
	Kaiwaiwai	0.17	16.51	0.86	83.49	1.03
	Paremata	0.47	35.72	0.85	64.28	1.32
	Rotopiko	3.74	24.81	11.33	75.19	15.07
	Taranaki	0.56	44.47	0.70	55.53	1.26
UAV	Horizons	0.46	87.14	0.07	12.86	0.53
	Kaiwaiwai	0.44	69.46	0.19	30.54	0.63
	Paremata	0.20	17.92	0.92	82.08	1.12
	Rotopiko	5.10	34.25	9.79	65.75	14.89
	Taranaki	0.46	38.75	0.73	61.25	1.19

These results provide actionable data on the riparian survival and ecological succession at these sites. Based on a calculation that compares the median intended % of total cover of all 5 pilot sites (Table 8) of each sensor, the Maxar, UAV and Aerial models all had a comparable consistency in predicting intended cover. The results of Paremata and Rotopiko (sites with the most detailed planting information) suggest very high ‘intended’ median cover percentages with Paremata at 87.15% and Rotopiko at 75.89%.

Table 8: Median intended % of total cover and nearest to median sensor

Site	Intended % S2	Intended % Maxar	Intended % Aerial	Intended % UAV	Median %	Nearest sensor
Horizons	7.65	88.97	13.69	12.86	13.275	Aerial/UAV
Kaiwaiwai	78.69	96.37	83.49	30.54	81.09	S2/ Aerial
Paremata	97.24	92.15	64.28	82.08	87.115	Maxar/UAV
Rotopiko	83.56	76.59	75.19	65.75	75.89	Maxar/Aerial
Taranaki	69.26	65.88	55.53	61.25	63.565	Maxar/UAV

Figure 6 presents a visual presentation of the output of all the models for each pilot site.

Based on the visual output for Rotopiko, the Sentinel 2 results are very coarse and primarily determining the unvegetated and grass classes accurately. The results of the Aerial model are generally very good and provided consistent results for most sites.

The Maxar model performed well for unvegetated and grass classes while slightly overpredicting of herbaceous and woody. The Maxar models overprediction of the woody class is a result of visual confusion with areas which should instead be herbaceous.

The visual output results of the UAV model indicate very good riparian class predictions, with the model being able to detect undesirable vegetation and distinguishing between sedges and herbaceous cover, although the model failed to detect sedges in the other pilot sites. Based on overall visual comparison and model F1 scores (Table 5) the UAV model provided the best detection.

The Maxar model presented the most consistent visual output compared to Aerial and Sentinel 2, and this corresponds with Maxar having the highest mean average accuracy compared with the field validation (Table 4). Table 9 provides a concise summary of the pros and cons of each model.

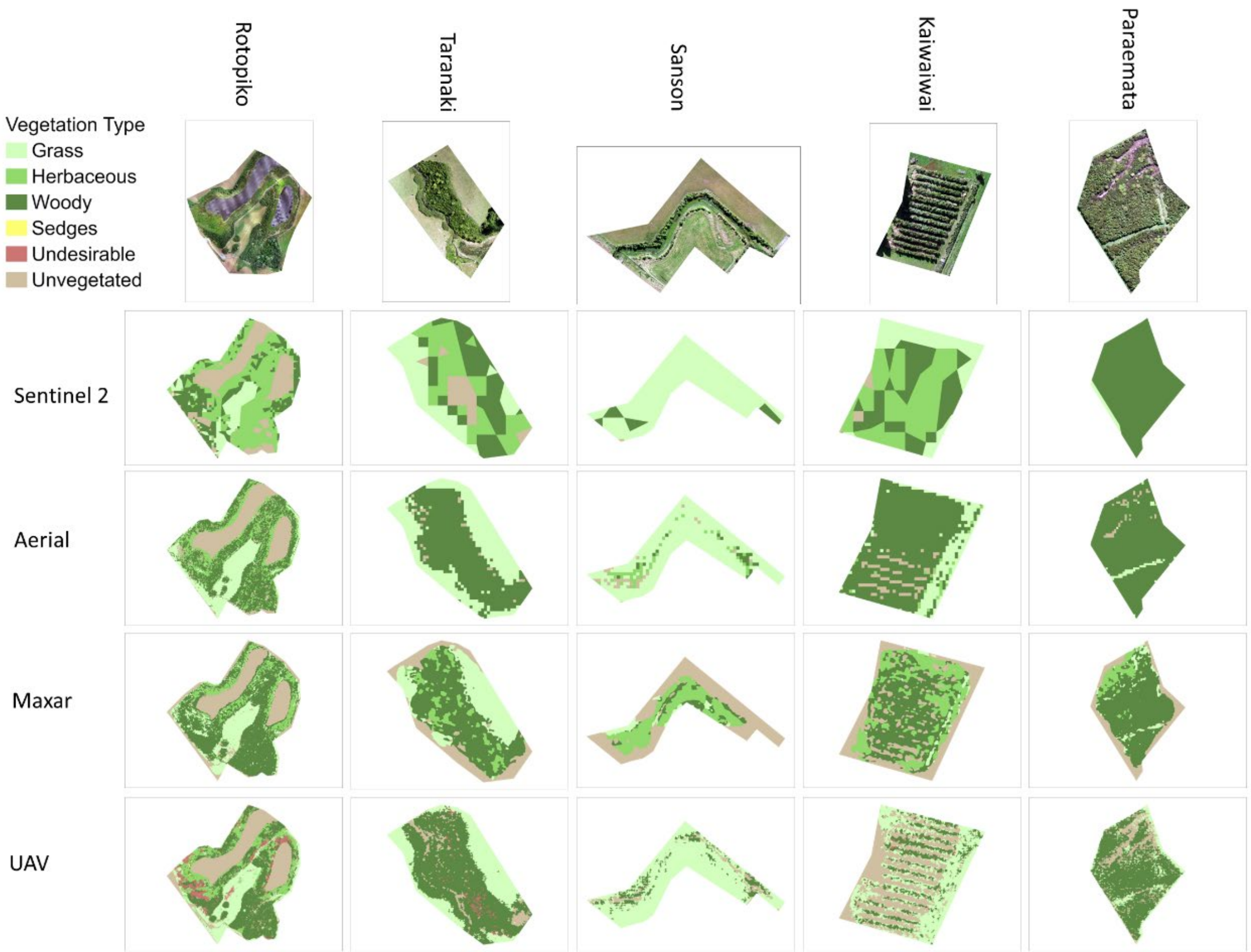


Figure 6: Model output results

Table 9: Model Pros and Cons

Sensor	Pros	Cons
S2	• Freely available • Good temporal and spectral resolution • Relatively small data	• Spatial resolution too low for machine learning training • Colour and texture not identifiable • Can't monitor areas of smaller spatial extents
Aerial	• Regional data freely available • Relatively good spatial resolution • Relative clear/sharp images for ML training of dominant vegetation patches (at 30cm) • Low spectral resolution	• Low temporal resolution (may be expensive and impractical for capture in winter season) • 30cm spatial resolution too low for training of sedges, undesirable and small patches of herbaceous vegetation • Expensive to acquire on demand
Maxar	• Cost effective for regional monitoring • Very good temporal resolution • Very good spectral resolution • Higher spatial resolution provides better ML training than S2 • Other options including Airbus available	• SecureWatch spatial resolution too low for precise RGB annotation ML training • Difficult to discriminate between vegetation colour and texture especially for the classes of grass, herbaceous and undesirable vegetation and even certain woody species
UAV	• Very high spatial resolution • Excellent ML training even undesirable herbaceous vegetation and sedges • Provide opportunity for high temporal resolution • May substitute field validation	• Logistics i.t.o CAA requirements and landowner permissions • Weather dependent • Standardisation of image capturing and processing that are adequate for ML training • Expensive to use for monitoring of regional wide plantings

2.3 Stage 1 Conclusions

The S2 model was too coarse for determination of riparian survival. The Aerial model performed well for most sites and provided consistent results although the low temporal resolution makes it impractical for assessment of riparian survival.

The UAV model provided the finest-scale mapping of riparian survival including unintended species although there was underprediction of herbaceous in some pilot sites. The UAV model represents an opportunity for high-value sites although the sensor is impractical to use once multiple sites need to be assessed due to the associated increased time and cost requirements compared to the alternatives. The timing of field validation work with the UAV acquisition additionally provided valuable insights into riparian survival seasonal constraints. Future studies should also include a comparison between summer and winter models to attempt to discriminate between riparian vegetation and weeds.

The Maxar model provided the most consistent outputs across the 5 pilot sites although this sensor includes some confusion between woody and herbaceous classes. We note that within the image resolution afforded by Maxar, it is not always easy to distinguish between low woody and herbaceous plants. Based on these results a national system that makes use of Maxar imagery using UAV validation data is very promising.

3.0 Stage 2 All Sites Analysis

This stage included the application and development of riparian planting survival plant classification models for at least 99 sites of MfE's choice using machine learning (ML) for the imagery analysis. A larger study area of 10x10 km in Taranaki was selected which included 141 sites (Figure 5) planted between 2017 and 2020.

At the commencement of stage 2 it was agreed with the Ministry to use a Maxar imagery model validating the model against botanical specialist desktop Maxar and UAV validation data.

3.1 Stage 2 Methods

The stage 2 analysis included the use of a 4-band Maxar input (Red, Green, Blue and Near Infrared) imagery for the ML model. Maxar 8-band imagery was used in the preparation of training datasets (referred to as the remote sensing model) but was not used by the ML model.

Validation data included 10 planting sites specifically held for botanical specialist validation (i.e., not used for ML) and model accuracy assessment. Eight of these sites were located within the 10x10 km Taranaki study area and 2 additional sites with UAV imagery coverage were selected - one in the Horizons region (Manawatū-Whanganui) and one north of the other sites in Taranaki (Figure 7).

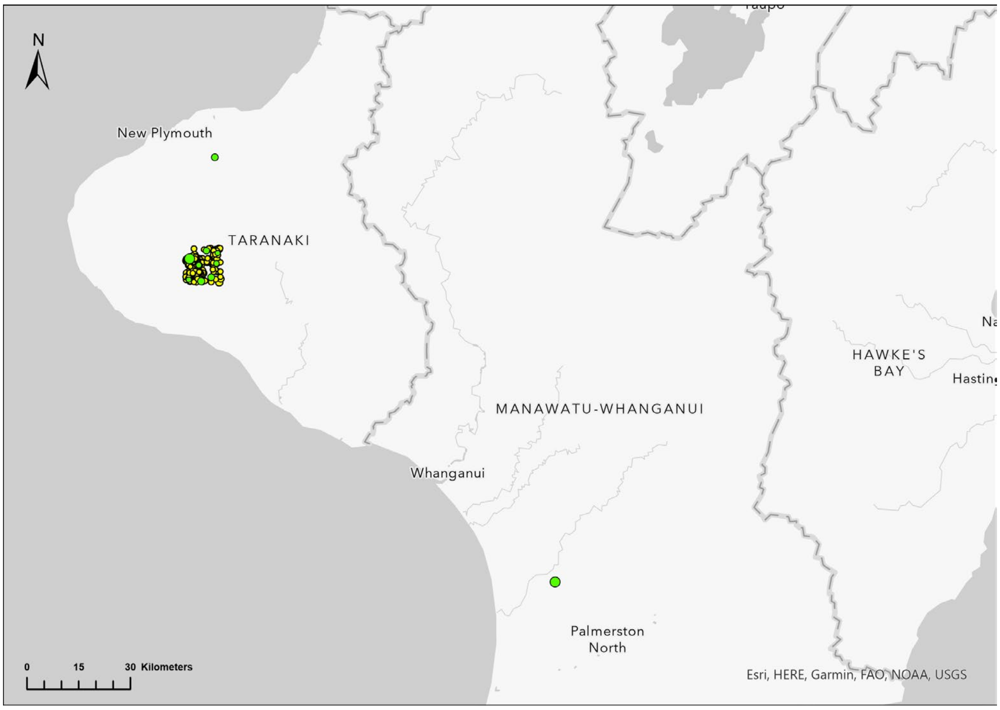


Figure 7: Stage 2 training sites (yellow) and validation sites (green).

Pan sharpening was performed on the 4-bands used for the ML model using the Gram-Schmidt spectral sharpening with Maxar (worldview) pre-set band weights. This process was performed in ArcGIS using the “Create Pansharpened Raster Dataset” tool. The inclusion of the Maxar multispectral bands(Maxar 8-band imagery) was explored and analysed to determine if these can improve plant classification for Stage 2 since only RGB imagery was used in Stage 1. Through the use of remote sensing methods (explained below) it was found that vegetation could be classified in more detail than the Stage 1 Maxar 4 class classification. This paved the way for a 6 class Maxar model for the Stage 2 analysis, which uses the following classes:

1. Tall Woody
2. Low-Medium Woody
3. Herbaceous
4. Rank Grass
5. Pasture Grass
6. Unvegetated

A 4-class classification was still used for accuracy assessment (validation) which consisted of dissolving the two woody and grass classes into just ‘woody’ and ‘grass’. Table 10 provides an overview of all the data inputs used for the Stage 2 analysis.

Table 10: Input Data Used in Stage 2 Analysis

Data Set	Description
Maxar SecureWatch	• Maxar panchromatic band (approximately 0.39 m) and spectral bands (B,G,Y,R, RE, NIR1) at 1.55 m resolution • 2021 and 2022 summer imagery • Sun elevation close to midday • Off-nadir (20-30 degrees) • <15% cloud cover • Coverage of a 10x10km all sites study area incl. 2 additional sites
Training data	• Human annotated training data • RS Model (RS training inputs) • 141 sites training sites • ⅓ randomly selected to measure model performance • ⅔ sites directly used by ML model
Validation data	• 10 validation sites for testing model accuracy • Botanical specialist desktop annotations using pansharpened Maxar imagery • 2 of these sites additionally annotated using high-resolution UAV
Planting data	• MfE and HRC planting data. • Planting dates ranged from 2017-2020 for the Taranaki study area. The planting age at the time of the Maxar image ranged from 4 months to 48 months.

For stage 2, we use a ML training methodology called supervised learning. This is essentially learning by example. By providing sufficient examples of the classified landcover output, paired with the Maxar input imagery, the machine learning model learns to produce a similar output when given the same imagery and, if trained well with sufficient data, will also produce classified landcover output when given new or previously unseen imagery as input.

A challenge in supervised learning projects is to find or create sufficient labelled data to train models well. We used a combination of two methods to prepare the training data (Figure 8). First, we generated a fully human annotated training dataset using super-pixels, then we used remote sensing methods using the multispectral bands from 8-band Maxar imagery to create a remote sensing training model. We then combined these datasets to create the full training dataset.

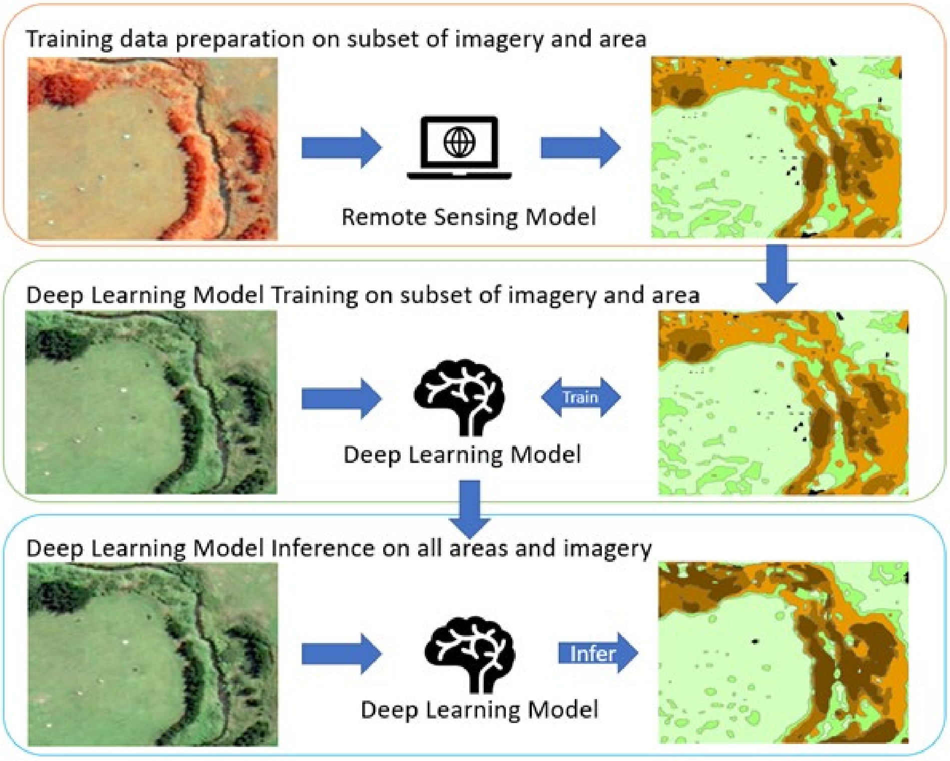


Figure 8: Model training flowchart

“Super-pixels” are a way of automatically segmenting an image. Super-pixels are clusters of neighbouring pixels within an image that look similar to each other, for example, they may have similar brightness or colour. There are many algorithms available to create super-pixels. Within the ESRI ArcGIS environment, there is Segment Mean Shift. This algorithm groups together adjacent pixels that have similar spectral and spatial characteristics. A human annotator can then assign a class to each of the segments.

The segments are prepared using two tools within ArcGIS; segment mean shift and raster to polygon. The output of this is a polygon layer that segments the image into areas of similarity spatially and spectrally. These tools are shown in Figure 9.

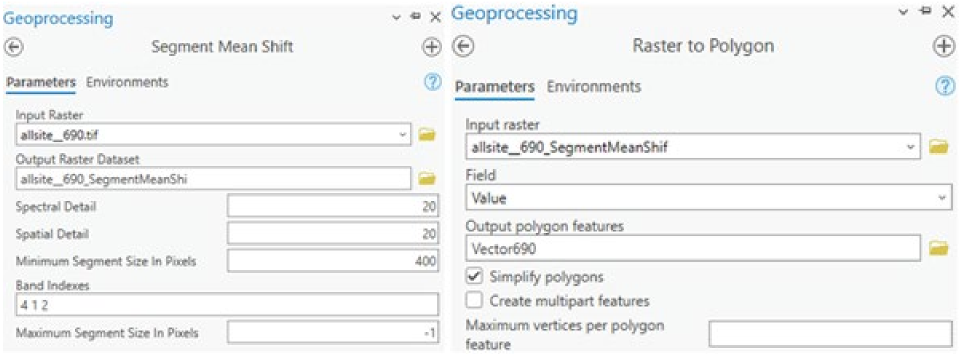


Figure 9: Super-pixel creation in ArcPro

The process of converting the image to a segmented image and then assigning each segment a class is shown in Figure 10.

Due to the high labour cost of this process, only two sites were manually classified in this way. These sites were left out of the subsequent training, monitoring and validation data sets. However, inference from this model was used as part of the training data for the final 6-class model as will be described further below.

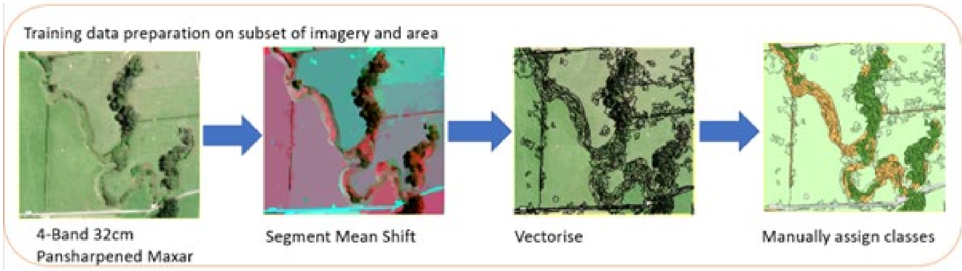


Figure 10: Super-pixel annotation of training data

The remote sensing model is prepared using a combination of spectral indices and a texture index. Vegetation indices are great for understanding if vegetation exists at a location and if the vegetated features are healthy, but they are not as effective at understanding the stage of growth, height, or density of plants reliably from a single high-resolution image. A texture measurement alone cannot identify vegetated areas from non-vegetated and a vegetation index cannot distinguish between short and tall vegetation. Using these in combination can help identify vegetative features and determine the density and height of vegetation (Figure 11). This remote sensing model is called the Vegetation Texture Index (VTI). Figure 12 presents a simplified overview of the VTI workflow. VTI is computed using the Maxar high resolution panchromatic band and lower resolution spectral bands Blue (B), Green (G), Yellow (Y), Red (R), RedEdge (RE) and Near Infrared (NIR1) depending on the vegetation index used. Pre-processing includes recoding of the panchromatic band and atmospheric correction of the multi-spectral bands. Texture is computed using r.texture (developer). The vegetation indexes used included Normalized Difference Red Edge index (NDRE) to estimate chlorophyll content in leaves, and the Red Edge Yellow index (REY) for the mapping of woody vegetation. The texture and vegetation index outputs are combined to form the VTI.

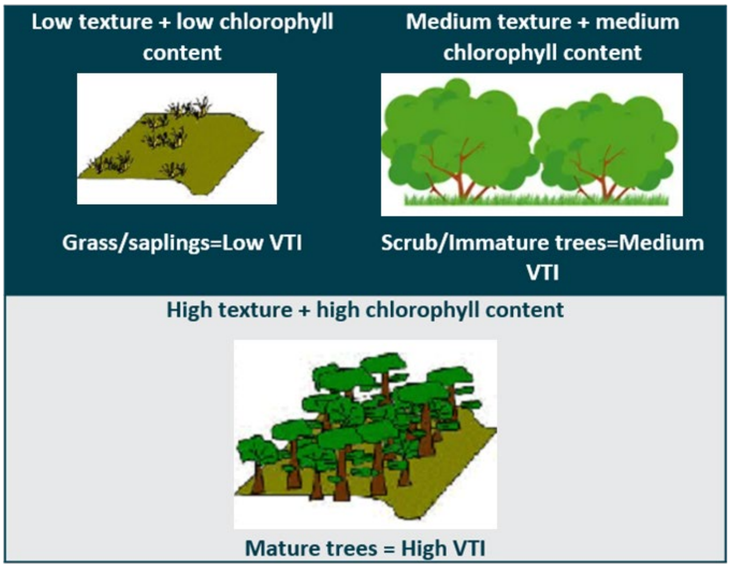


Figure 11: Vegetation Texture Index (VTI) Analysis (adapted from Maxar, 2022)

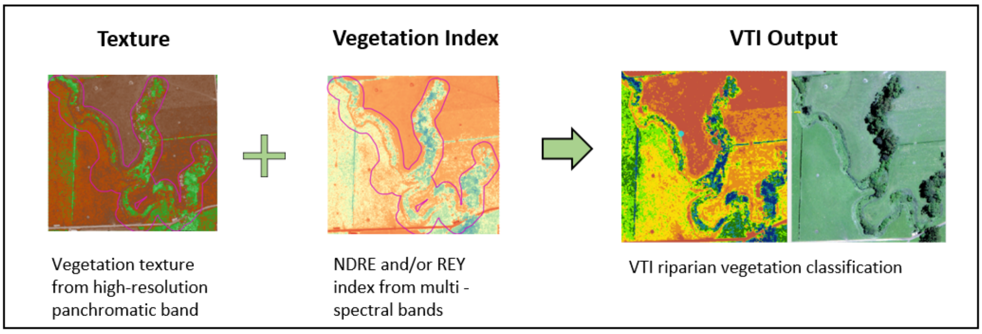


Figure 12: VTI workflow

A combination of several outputs was used to create the training data for the machine learning model. For example, we found that the NDRE VTI layer was more sensitive to the "tall woody" class and the 4-class machine learning model trained using human annotations of super-pixels was good at predicting grass. Each of the VTI outputs (NDRE and REY) were classified to create 6-class landcover layers. These were combined with the outputs of a model trained on the 4-class super-pixel hand labelled data (Figure 8 and Figure 13), and an NDVI with threshold to boost the detection of the unvegetated class. All training – including training of prior models was done on the training sites to keep independence from the validation sites.

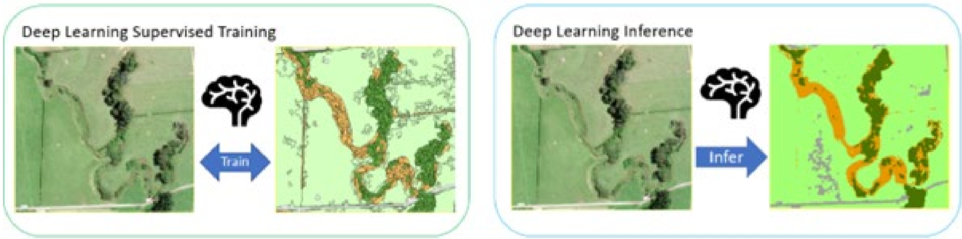


Figure 13: Super-pixel training of deep learning model

To train the machine learning models, a random one third of the sites are set aside from the training data and are used to measure model performance while training (the "monitoring" dataset). As a random subset of the training-monitoring data, this should have similar characteristics to the training data. The remaining training data is used directly by the machine learning model during model training, and it is from this data that the model learns the mapping from input 4-band imagery to output landcover classes.

The training and monitoring data is prepared in the image and classified landcover pairs of raster chips with size 128x128 pixels. The input imagery is 4-band Maxar and the output landcover has 6 classes.

During inference, overlapping chips of 128x128 with a 64 pixel overlap with neighbours on each side are used and merged in weighted fashion to create a smooth inference across the whole input raster. The input raster can be of arbitrary rectangular size – within computational limits.

Stage 2 validation

Ten test sites were held out to evaluate the final model accuracy (Figure 5 and Table 9). These sites were not included in the training or monitoring datasets. A desktop analysis was performed by a human expert to identify points and classes within these sites as "Ground Truth" against which the model performance was measured. We used a dedicated validation data platform on ArcGIS online (AGOL) with the plant classification schema for the creation of desktop validation data (Figure 14). Pansharpened Maxar imagery acquired of the Taranaki study area and the two sites falling outside of the study area served as the main input layer for the validation. High-resolution (0.03m) RGB UAV imagery was additionally also used for validation (Figure 14) for the 2 sites outside of the Taranaki study area.

The desktop analysis attempted to label representative points and to have ten points per class per site. However, the class presence on the ground made this difficult as some classes such as "herbaceous" are not evident at all sites. Similarly, unvegetated areas were prominent in some sites but not in others. We also note that the objectives of representative sampling and sampling equal numbers per class are inconsistent with each other as actual class distribution is unbalanced with woody and grass classes dominating at most sites.

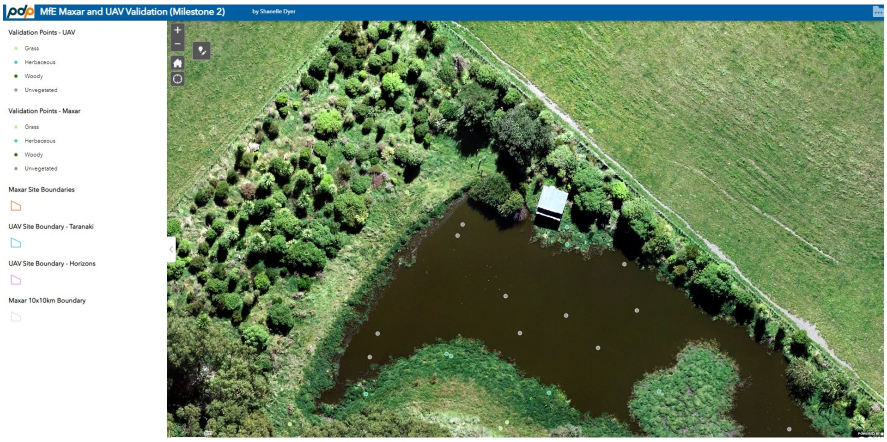


Figure 14: Screenshot of the Desktop validation platform using UAV imagery

<u>Machine Learning Model</u>

We use a machine learning model architecture called DeepLabV3+ which is a well-known and efficient convolutional neural network for semantic segmentation (Figure 15). DeepLabV3+ was developed by researchers from google in 2018 and despite its great age in deep learning terms, is still high performing on standard benchmarks.

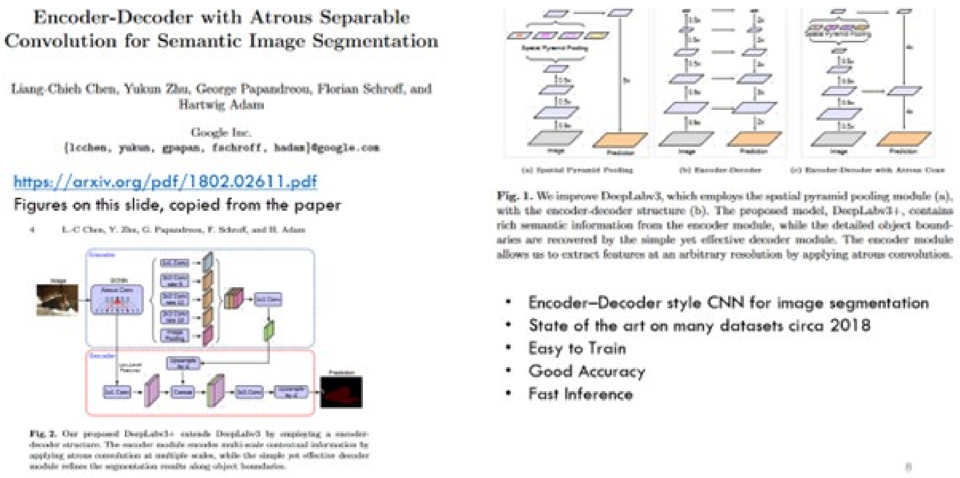


Figure 15: DeeplabV3+ overview

DeepLabV3+ takes in RGB images and in a normal supervised learning task will learn to output a segmentation mask (Figure 16) once trained on thousands of example image and segmentation mask pairs.

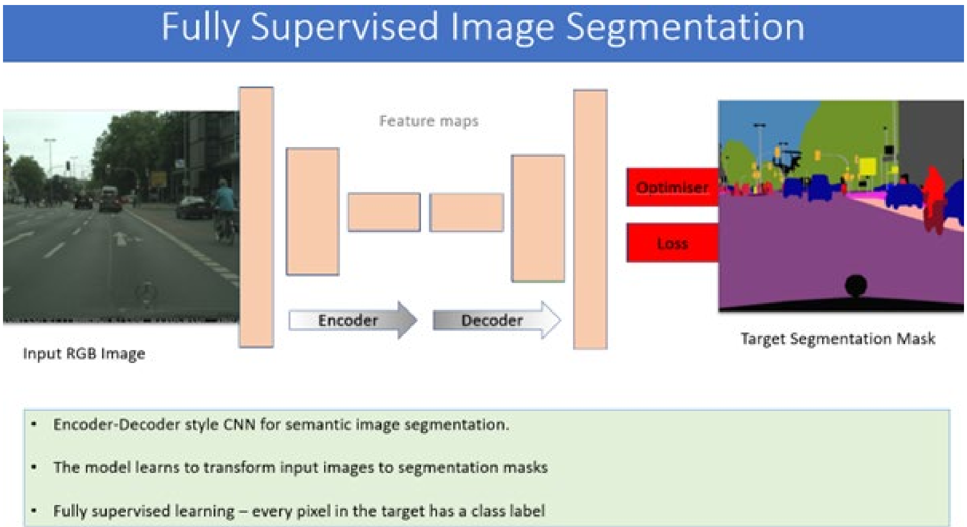


Figure 16: Example of supervised learning, DeeplabV3+

We modified the input to standard DeepLabV3+ to accept 4 band input instead of the standard 3 band input.

Our implementation of DeepLabV3+ uses a ResNet50 backbone. ResNet50 is an image classification model and has pre-trained weights trained on the ImageNet dataset. Using a pre-trained model means that the model already "knows" abstract representations of information within an image. Because of this prior knowledge, the model should learn new classifications more efficiently – with less training data required to achieve reasonable accuracy. We train the model using a method called “finetuning”.

Finetuning broadly means starting with an already trained model and training it on a new dataset. In DeepLabV3+, the encoder part of the network is pre-trained on ImageNet, but the decoder part of the network has not been pre-trained. We begin training using the training dataset to train the decoder part of the network only. Then, we reduce the learning rate of the network and train all parts of the network until the accuracy, as measured against the monitoring dataset, stops improving.

To modify the network to accept 4-band input while still using a pre-trained backbone (that accepts only 3-band RGB imagery), we create a small pre-network that accepts 4-bands and learns how to condense these to 3-bands. We found through experiments that this method improved training and monitoring accuracy by a few per cent over training and inference using 3-band RGB imagery alone.

3.2 Stage 2 Results

Model accuracy was assessed against the validation data. The validation sites had not been seen by the model prior to the final inference.

<u>Taranaki study area validation</u>

Table 11 shows the classification accuracy summary across the 4-class classification and the validation sites within the Taranaki study area. The overall f1-score and classification accuracy are both 0.72 (72%). The “Support” column shows the number of labelled ground truth points with the class indicated by the row.

We note the relatively high f1-scores for Grass and Woody classes and the relatively low f1-score for herbaceous. We speculate that this is due to the relative dominance of grass and woody classes at all sites and so reflects the amount of training data available for each class. Please refer to **Appendix A** for a classification accuracy for each individual site including the accuracy against the UAV validation data.

Table 11: All Sites classification accuracy

	Precision	Recall	F1-Score	Support
Grass	0.72	0.89	0.8	80
Herbaceous (Herb)	0.39	0.45	0.42	31
Unvegetated (Unveg)	0.98	0.53	0.69	81
Woody	0.72	0.85	0.78	78
Accuracy			0.72	270
Weighted average	0.76	0.72	0.72	270

Figure 17 presents a confusion matrix; it shows the ground truth class in the rows (with headings on the y-axis) and the predicted class in columns (with headings on the x-axis). We see good agreement between ground truth and predicted class for Woody and Grass classes but poor agreement for Herbaceous. We note that true herbaceous of often predicted to be Grass. We also note that true Unvegetated is sometimes predicted to be Woody. Some examples of why this might be are discussed later in this section.

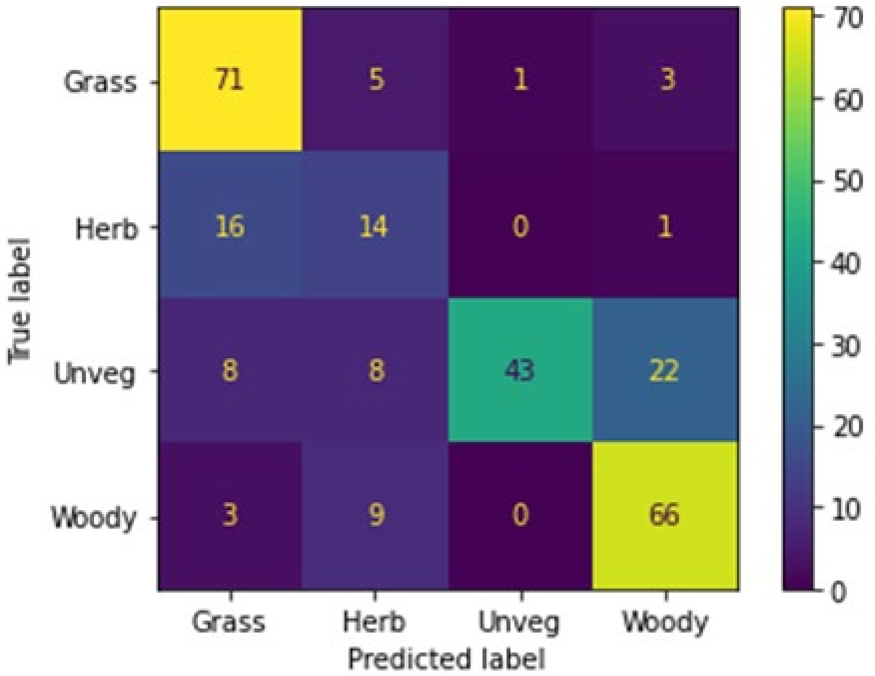


Figure 17: Taranaki study area validation sites classification confusion matrix

In addition to these eight sites, we tested the model on the two additional validation sites. Each used a different source scene from Maxar including validation against UAV imagery.

Taranaki Additional Validation Site

We note the lower overall accuracy (Table 12) when compared to the 8 sites within Taranaki study area. Classification accuracy is 0.62 and weighted average f1-score is 0.59. We observe a similar accuracy pattern across classes with a low f1-score of 0.13 for Herbaceous. We examine the classification confusion matrix, shown in Figure 18 and observe perfect recall for Woody and good agreement for Unvegetated but observe that true Herbaceous is often predicted to be Woody.

Table 12: Classification accuracy for Taranaki additional validation site				
	Precision	**Recall**	**F1-Score**	**Support**
Grass	0.75	0.6	0.67	10
Herbaceous (Herb)	0.2	0.1	0.13	10
Unvegetated (Unveg)	0.53	1	0.69	10
Woody	1	0.8	0.89	10
Accuracy			0.62	40
Weighted average	0.62	0.62	0.59	40

Figure 18: Confusion matrix Taranaki additional validation site

UAV imagery was used in addition to the Maxar imagery for validation over the Taranaki additional site. This imagery was co-registered with the Maxar imagery. Desktop validation points were chosen in the UAV imagery and labelled by a botanical specialist. The accuracy of the model was then evaluated against this data. The model and model inference were the same but the validation points were different to those evaluated above. With this approach we see poor accuracy for the unvegetated and herbaceous classes. We consider that this is a consequence of the fine spatial detail and the unvegetated and herbaceous classes being represented by small spatial features that were below the resolution of the model. The overall classification accuracy against the UAV data was 49% (Appendix A). This accuracy is affected by co-registration, timing of imagery and detail available from each sensor.

Horizons Validation Site

We observe a classification accuracy of 0.65 and a weighted mean f1-score of 0.6 (Table 13). We observe the same pattern of low accuracy for the Herbaceous class.

The confusion matrix (Figure 19) shows a good agreement between ground truth and model for the Unvegetated, Woody and Grass classes and low agreement for Herbaceous. We observe that true Herbaceous is often predicted to be Grass and that the model underpredicts Herbaceous.

Table 13: Classification accuracy for Horizons

	Precision	Recall	F1-Score	Support
Grass	0.53	0.8	0.64	10
Herbaceous (Herb)	0.33	0.1	0.15	10
Unvegetated (Unveg)	0.78	0.7	0.74	10
Woody	0.77	1	0.87	10
Accuracy			0.65	40
Weighted average	0.6	0.65	0.6	40

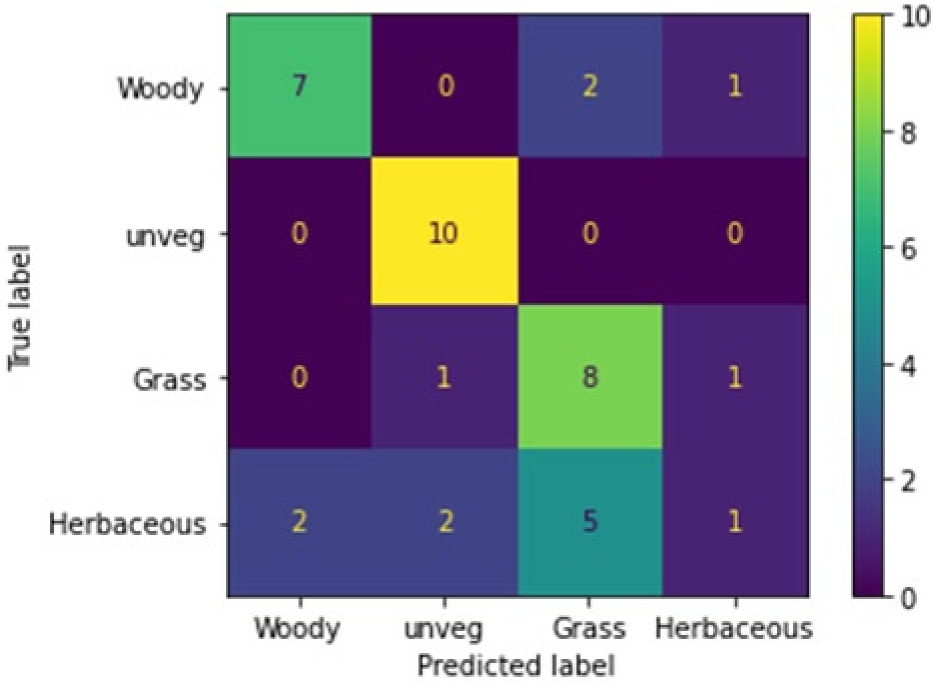
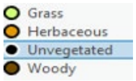
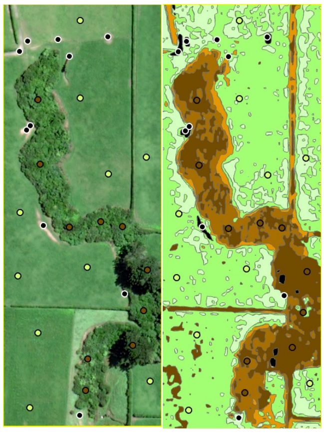
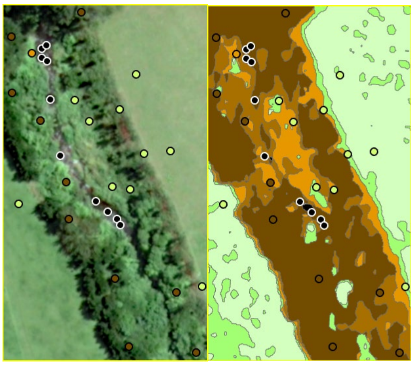


Figure 19: Confusion matrix Horizons validation site

Additional data validation points were gathered from UAV imagery in the same way as for the Taranaki additional validation site. As for Taranaki, we see lower accuracy of the model against this validation data. The classification accuracy for the Maxar model against the UAV gathered validation data is 58% (Appendix A).

Table 14 presents a visual comparison of two selected sites. The first site provides an example of a good accuracy and the second with a low accuracy where the unvegetated narrow waterway detail is lost.

Table 14: Visual comparison with validation data

Site	Legend	Model output with validation	Finding
Taranaki ID 1318	**Validation**  **Model**  		Site with good accuracy: 90%
Taranaki ID: 4595			Site with low accuracy:49% Narrow waterway detail lost which resulted in a low unvegetated class accuracy. In future we recommend watercourses are masked prior.

Notes:

1. *Model output overlayed with validation: left Maxar RGB right 6-class inference.*

<u>Masking known waterways</u>

We observe from Table 15 that the model differs most from the validation points where we have unvegetated validation points in narrow river channels. In phase 1 of the project, we worked with the assumption that known waterways would be masked out of the analysis. Applying that same assumption to the phase 2 validation and excluding the unvegetated points within waterways improves the measured model performance to 80%. The following table shows the overall model accuracy where river or stream water has been masked out.

Table 15: All Sites classification accuracy with masked waterways

	Precision	Recall	F1-Score	Support
Grass	0.72	0.89	0.8	80
Herbaceous (Herb)	0.48	0.45	0.47	31
Unvegetated (Unveg)	0.98	0.8	0.88	54
Woody	0.92	0.85	0.88	78
Accuracy			0.8	243
Weighted average	0.81	0.8	0.8	243

3.3 Determination of planting success from model output

Riparian zones are defined as "strips of land beside drains, streams, rivers and lakes." They are priority locations for planting native vegetation to maintain bank stability, reduce sediment run-off, shade waterways, and provide habitat for indigenous species. They are often planted as narrow bands of specific vegetation types (or classes), including a 1 m strip of land often intentionally left unplanted to develop into rank grass as a filter and to protect stock fences (See Figure 20). Upper banks are often planted in taller, woody vegetation (trees, shrubs, cabbage trees, and treeferns), while steeper lower banks are often planted in flood-tolerant herbaceous vegetation (sedges, flaxes, and tussocks). True indigenous grasses are rarely planted in New Zealand riparian projects.

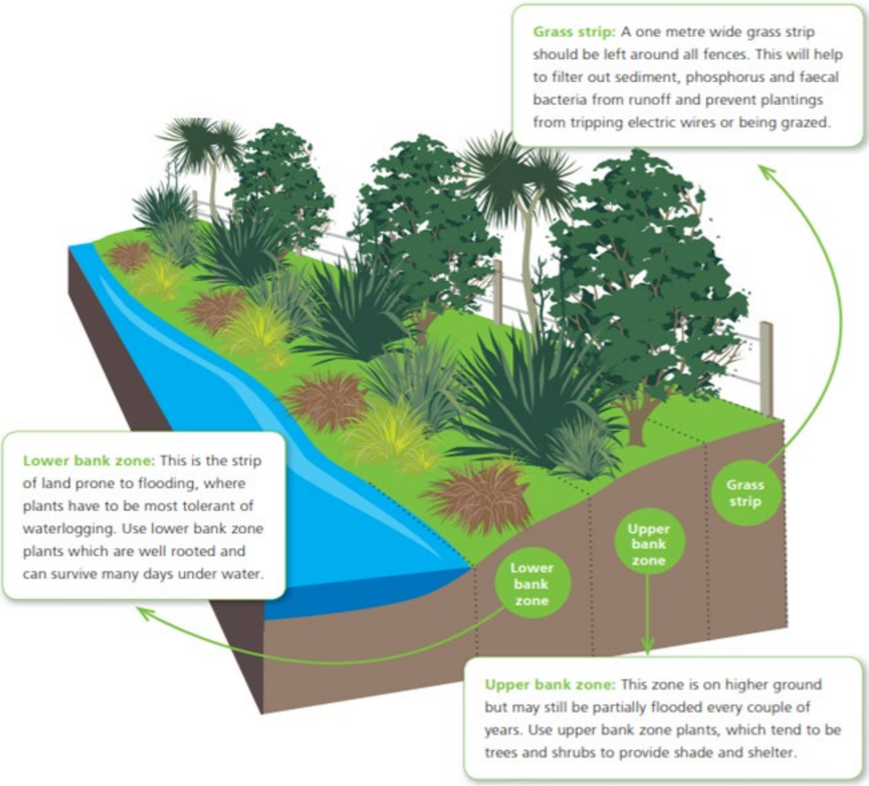


Figure 20: Schematic of typical riparian zones (adapted from Dairy NZ, 2014).

The Taranaki study area planting sites included a large variation in planting method and all the sites used in stage 2 was planted mostly with woody plants. The stage 2 model results indicated that herbaceous plants was detected with a much lower precision compared to woody plants. Woody vegetation coverage therefore serves as the main indicator for planting success in the current model (stage 2 model). Woody classes in the model included tall woody and low-medium woody which make up woody. The current model does not distinguish between desirable and undesirable woody. Undesirable woody species may include species such as woolly nightshade, privet, willow, broom and gorse. Woody vegetation shade water waterways and provide habitat for indigenous species. When a woody coverage (canopy closure) of 80-100% is reached it shades out and excludes weeds.

Coverage (canopy closure) is based here on >1 m to <5 m plant spacing. Coverage included either those deliberately planted (intended species) or self-seeded native plants that are ecologically appropriate to the local area and riparian zone. This assessment allows for undesirable/unintended species to comprise up to 10% of that portion of the PPZ no longer in residual grassland/bare ground. Example if 80% of the site has a cover of woody vegetation, the target is met if 72% of the total planting zone cover is intended vegetation, and 8% is unintended vegetation (the balance 20% being residual grass/bare ground).

Table 16 present the steps to determine planting success from the stage 2 model output. Please start with the column on the left and follow the footnote references.

Table 16: Planting success from image parameters

IF Target Vegetation =	AND Atkinson* classes =	AND Months between planting and image date =	AND Coverage[4] (%) =	Expected Height (Range stature) =	Survival = (Acceptable or not yet acceptable)
Woody[1]	Forest[2] or Scrub	>=60	> 60	Medium or tall	Acceptable
Woody[1]	Treeland[3] or shrubland	>24 <60	> 30	Medium or tall[5]	Acceptable
Woody[1]	Treeland or shrubland	<=24	> 10	Low[6]	Acceptable

Notes:

1. *Woody vegetation is the main indicator for planting success in the current model. Woody classes in the model included tall woody and low-medium woody which make up woody. Woody species dominated the plantings in the planned planting zone (PPZ). The current model does not distinguish between desirable and undesirable woody plants. Undesirable woody species may include species such as woolly nightshade, privet, willow, broom and gorse. Woody vegetation shade water waterways and provide habitat for indigenous species and when a coverage (canopy closure) of 80-100% are reached it shades out and excluded weeds (unintended species).*
2. *It is unlikely that planted vegetation would meet the definition of forest or treeland (woody plants >10 cm trunk diameter) within the 5-year time frame of the riparian assessment, however forest and treeland are listed here for completeness.*
3. *As above.*
4. *Coverage (canopy closure) based here on >1<5 m plant spacing. Coverage here includes either those deliberately planted (intended species) or self-seeded native plants that are ecologically appropriate to the local area and riparian zone. This proposal allows for undesirable/unintended species to comprise up to 10% of that portion of the PPZ no longer in residual grassland/bare ground. Example if 80% of the site has a cover of woody vegetation, the target is met if 72% of the total planting zone cover is intended vegetation, and 8% is unintended vegetation (the balance 20% being residual grass/bare ground).*
5. *Tall vegetation in the >24 <60 months category indicates a high percentage of existing woody cover not part of the planting in the PPZ. Coverage (canopy closure) are ranked as the main measure of planting success .since there is a high variability among species, in terms of height. A high % of tall woody indicates the maturity of the vegetation. A >=30% cover of desirable tall woody in the >24 <60 months category is favarouble in terms of ecological succession.*
6. *The riparian model output classifies grass and herbaceous as vegetation with a low range stature. A sporadic woody cover with a planting age of <=24 months have a similar signature i.t.o plant stature as grass and herbaceous. The riparian model therefore does not distinguish between these classes i.t.o plant stature.*
7. *The grass classes likely represent failed planting as true native sward-forming grasses are unlikely to be planted in NZ riparian projects, however narrow strips of rank grass are often an intended outcome to protect fences and trap silt.*
8. *The herbaceous class in the current model do not distinguish between desirable and undesirable plants. The herbaceous cover in the study includes mainly self-seeded plants. This class may include native plants that are ecologically appropriate, weeds including sporadic woody cover (see note 5).*
9. ** Atkinson 1985 vegetation structure classes*

Table 17 and Figure 21 presents an example of a riparian survival assessment for site (global ID: BECADBDB-58CA-451D-9F43-CDA379CD01DD) using the stage 2 model output. The targeted number of plants for this site was 110. The planting date was 15/08/2019. The age of the planting was 17.5 months at the time of the Maxar image. The PPZ was estimated for the testing the determination of planting success from model output as there is no available info on the actual planting zones. The total woody cover was 45.94% with 27% of the woody cover consisting of Tall Woody which indicate that the site already included a large woody cover at the time of planting. This further indicate the maturity of the woody cover which is nearing 30% just for the Tall Woody. The site (PPZ) included a 22.7% rank grass coverage which may indicate areas of failed planting, areas not planted, young sporadic woody cover/woody with low foliage diameter. This assessment results indicate an acceptable survival since the woody cover are 45.94% since it needs to be over 10% for the under 24 months category. The high percentage medium to tall woody vegetation ad confidence including having a good detection of areas covered by grass.

Table 17: Model results of a specific site

Predicted6[1]	**RangeStat[2]**	**Predicted4[3]**	**Percentage cover[4]**	**Area (ha)[5]**
Tall Woody	Tall	Woody	27.12	8.13
Low-Medium Woody	Medium	Woody	18.83	5.65
Herbaceous	Low	Herbaceous	25.52	7.65
Rank Grass	Low	Grass	22.73	6.82
Pasture Grass	Low	Grass	5.73	1.72
Unvegetated	Unvegetated	Unvegetated	0.071	0.021
			100	**29.98**

Notes:

1. *Predicted6: 6-class model output*
2. *RangeStat (plant stature) Low=grass & herbaceous, Medium=low -medium woody, High =tall woody*
3. *Predicted4: 4-class model output*
4. *Percentage cover: % cover per class per site*
5. *Area (ha): area in hectares per vegetation class*

Figure 21: Example of riparian model output for a specific site

Determination of planting success using winter models

The utilisation of winter Maxar on its own or in combination with summer imagery may provide the opportunity to utilise a more detailed multi-variate test of riparian planting success compared to the test in Table 16 that focus on woody vegetation as the main indicator of planting success. Table 18 presents an expanded test that includes herbaceous (includes sedges) and grass classes for the determination of planting success.

Table 18: Planting success from image parameters

IF Target Vegetation =	AND Atkinson* classes =	AND Months between planting and image date =	AND Coverage[3] (%) =	Expected Height (Range stature[4]) =	Survival = (Acceptable or not yet acceptable)
Woody	Forest[1] or Scrub	>=60	> 60	Medium or tall	Acceptable
Woody	Treeland[2] or shrubland	>24 <60	> 30	Medium or tall[5]	Acceptable
Woody	Treeland or shrubland	<=24	> 10	Low[6]	Acceptable
Herbaceous/ Sedgeland	Treeland or Shrubland or Herbaceous	>=60	>60	Medium or Low	Acceptable
Herbaceous/ Sedgeland	Treeland or Shrubland or Herbaceous	>24 <60	>30	Low	Acceptable
Herbaceous/ Sedgeland	Treeland or Shrubland or Herbaceous	<=24	>10	Low	Acceptable
Grassland[7]	Any	Any	Any	Low	Acceptable[7]

Notes:

1. *It is unlikely that planted vegetation would meet the definition of forest or treeland (woody plants >10 cm trunk diameter) within the 5-year time frame of the riparian assessment, however forest and treeland are listed here for completeness.*
2. *As above.*
3. *Coverage (canopy closure) based here on >1<5 m plant spacing. Coverage here includes either those deliberately planted (intended species) or self-seeded native plants that are ecologically appropriate to the local area and riparian zone. This proposal allows for undesirable/unintended species to comprise up to 10% of that portion of the PPZ no longer in residual grassland/bare ground. Example if 80% of the site has a cover of woody vegetation, the target is met if 72% of the total planting zone cover is intended vegetation, and 8% is unintended vegetation (the balance 20% being residual grass/bare ground).*
4. *RangeStat (plant stature) Low=grass & herbaceous, Medium=low -medium woody, High =tall woody*
5. *Tall vegetation in the >24 <60 months category indicates a high percentage of existing woody cover not part of the planting in the PPZ. Coverage (canopy closure) are ranked as the main measure of planting success .since there is a high variability among species, in terms of height. A high % of tall woody indicates the maturity of the vegetation. A >=30% cover of desirable tall woody in the >24 <60 months category is favarouble in terms of ecological succession.*
6. *The riparian model output classifies sporadic grass and herbaceous as vegetation with a low range stature. A sporadic woody cover with a planting age of <=24 months have a similar signature i.t.o plant stature as grass and herbaceous. The riparian model therefore does not distinguish between these classes i.t.o plant stature.*
7. *The grass classes likely represent failed planting as true native sward-forming grasses are unlikely to be planted in NZ riparian projects, however narrow strips of rank grass are often an intended outcome to protect fences and trap silt.*
8. *The herbaceous class in the current model do not distinguish between desirable and undesirable plants. The herbaceous cover in the study includes mainly self-seeded plants. This class may include native plants that are ecologically appropriate, weeds including sporadic woody cover (see note 5).*
9. ** Atkinson 1985 vegetation structure classes*

4.0 Conclusion

The results produced in this assessment and the learnings derived from the modelling, training and tuning processes has yielded the following key findings.

1. Maxar imagery could effectively be used with the riparian model developed for this project, to determine riparian survival. The final prediction model reported an overall accuracy of **72%** and by masking known waters, the accuracy rises to 80% with unvegetated, woody and grass classes having F1 scores between 0.8 and 0.9. The model had difficulty resolving small narrow features such as narrow rivers that adversely affected the validation accuracy for the unvegetated class before masking known waterways. Certain sites which had very narrow features such as streams and an excess of validation points within those features had lower accuracy as a result prior to masking but with masking had improved accuracy. The model struggles with features that are smaller than 4 pixels and with pan-sharpened imagery, this can be slightly worse as the colour information comes from sensors that have lower resolution than the final pan-sharpened image. We see this problem more in narrow streams and narrow herbaceous strips. However, the overall landcover representation was good.

2. The developed model achieves the project objectives of determining riparian planting survival with an acceptable accuracy mapping riparian vegetation into 6 classes. The method can be upscaled to predict riparian planting survival for national programmes.

3. Although the spatial resolution of the Maxar worldview is high, it is still too low to distinguish between undesirable and desirable (unintended and intended) vegetation and there was still a degree of confusion between woody and herbaceous. It is recommended that future riparian survival analysis using this method should utilise winter Maxar imagery in combination with summer imagery where practical. Findings of this study suggests that winter imagery would be better suited for discrimination of herbaceous from woody plants including discriminating between riparian vegetation and weeds (undesirable vegetation).

4. The UAV model provides good opportunity for high-value sites although the sensor is impractical to use once multiple sites need to be assessed due to the associated increased time and cost requirements compared to alternatives.

5. The absence of detailed planting information especially the exact locations of the planned planting zones (PPZs) make the determination of riparian planting survival difficult. Without these it is not possible to distinguish between where planting occurred and existing vegetation. These boundaries are important for the calculation of planted riparian

coverage and ultimately riparian survival, to understand ecological succession and growth of self-seeded vegetation. It would greatly enhance assessment if applicants were required to supply the Ministry with planting plans in GIS format that clearly delineate the extents (footprints) of planting to ensure that existing vegetation are being excluded from riparian survival analysis.

In conclusion, the developed model achieves the project objectives of determining riparian planting survival with an acceptable accuracy, mapping riparian vegetation into 6 classes. The method can be applied to predict riparian planting survival across New Zealand.

5.0 Stage 2 Deliverables

Data delivery was completed in accordance with MfE requirements and was delivered as an Esri File geodatabase in NZTM 2000 projection.

The main data outputs include the stage 2 model riparian vegetation classification and related attributes including coverage of each class and the plant stature classification. The output includes a feature class with detailed planting information of each site. This includes estimated planting site boundaries including existing MfE and HRC planting information, for example Site ID, planting year and planting date. The image planting age (planting age at the time of the image) was added to facilitate the calculation of riparian planting survival from the stage 2 model output results.

6.0 References

Atkinson, I.A., 1985. Derivation of vegetation mapping units for an ecological survey of Tongariro National North Island, New Zealand. New Zealand Journal of Botany, 23, 361-378.Dairy NZ, 2014.

https://www.dairynz.co.nz/media/1653948/Riparian_Management_Canterbury_b.pdf

Denyer, K., 2021. MfE Riparian Planting Survival Assessment: Proposed Assessment Method and Classification Attributes. Unpublished project input report

Maxar, 2022. https://blog.maxar.com/tech-and-tradecraft/2022/texveg-the-next-generation-vegetation-index

Appendix A: Additional Tables

Table 19: Stage 2 classification accuracy for each validation site

Site	Sensor	Unvegetated F1	Herbaceous F1	Grass F1	Woody F1	Classification Accuracy
1318	Maxar	0.89	na	0.86	0.95	0.9
1336	Maxar	1	0	0.8	0.84	0.87
1359	Maxar	0.84	0.67	0.82	0.95	0.83
2303	Maxar	0	0.29	0.82	0.5	0.48
3206	Maxar	0.57	0	0.91	0.86	0.77
3391	Maxar	0.46	0.25	0.86	0.77	0.68
4595	Maxar	0.31	0.29	0.67	0.57	0.49
9912	Maxar	0.89	0.67	0.71	0.93	0.79
All-Sites summary	Maxar	0.69	0.42	0.8	0.78	0.72
Taranaki Additional	Maxar	0.89	0.13	0.67	0.69	0.62
Horizons Additional	Maxar	0.87	0.15	0.64	0.74	0.65
Taranaki Additional	UAV	0.29	0.14	0.9	0.47	0.49
Horizons Additional	UAV	0.87	0.12	0.43	0.75	0.58

Model output on an approximately 10km x 10km area of Taranaki including the training and validation sites, as shown in red.