Deforestation Mapping 2017 and 2018

Technical Report

Prepared for

Ministry for the Environment

13 November 2020

CITE REPORT AS:

New Zealand Deforestation Mapping 2017 & 2018 – Technical Report

Submitted by Lynker Analytics Consortium for the Ministry for the Environment

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning
AGL	Above Ground Level
AL	Active Learning
CAA	Civil Aviation Authority
CNN	Convolutional Neural Network
FMS	Flight Management System
FN	False Negative
FP	False Positive
GIS	Geographic Information System
GPS	Global Positioning System
GSD	Ground Sample Distance
LUCAS	Land Use and Carbon Analysis System
MfE	Ministry for the Environment
ML	Machine Learning
MPI	Ministry of Primary Industries
QAQC	Quality Assurance, Quality Control
RGB	Red Green Blue
TP	True Positive
TN	True Negative

Executive Summary

The Land Use Carbon Analysis System (LUCAS) is used for reporting on Land-Use Change and Forestry in NZ's annual national greenhouse gas inventory report submitted to the United Nations to meet New Zealand's international reporting obligations. From a greenhouse gas point-of-view, the most important land use changes are afforestation (new forest planting) and deforestation (forest removal and conversion to another land use).

The Lynker Analytics Consortium consisting of three companies; Lynker Analytics, Carbon Forest Services and UAV Mapping NZ Ltd were commissioned in late-2019 to survey and classify a total of 7,484 distinct areas of potential forest loss.

Using satellite imagery, the Ministry for the Environment (the Ministry) had previously mapped approximately 4990 distinct areas of forest loss, which occurred during 2017 and 2018. Added to this were 2494 areas that lost forest in earlier years, for which current land use was unknown. These 7484 forest-loss 'targets' range in size from approximately 1 to 500 hectares and cover a total area of approximately 83,600 hectares.

Between January and August 2020, the Lynker Analytics Consortium conducted an aerial survey of all areas spanning every region of New Zealand using Cessna 172 aircraft flying at approximately 5,000 feet above ground level (AGL). This delivered 0.25m resolution vertical aerial photography in over 99% of targets. A further four targets were acquired using ground photography while 11 were passed in for future survey (due to weather).


The imagery was georeferenced and then classified into land cover classes such as cutover, plantation seedlings, pasture, and mature native forest within a target using a Machine Learning (ML) approach. From this we applied a geospatial data generalisation routine to filter out noise or speckle and then we used a multi-criteria iterative analysis to assign each area of forest loss a dominant land cover and replant status.

The ML model was used as the primary monitoring system to detect deforestation and re-planting and flag those targets to the Ministry. Overall this approach allowed us to map land-use within target perimeters at a resolution of 100m^2 enabling greater alignment with the definitions of 'forest land', 'non-forest land', and 'deforestation' under the international and domestic rules.

The automated monitoring system proved reliable in detecting deforestation, re-planting and other land cover changes exceeding one hectare. It also enabled more rapid assessment of replant status used by the Ministry for reporting. In total 7473 targets were surveyed with imagery, land cover polygons, dominant land cover and replant status attributes provided to the Ministry.

1 Scope and Methodology

The Ministry undertakes mapping of deforestation every two years by first identifying areas of forest loss in satellite imagery and then determining which areas correspond to deforestation, as opposed to harvest activity, or forest loss due to natural causes. Over the summers of 2016/17, 2017/18 and 2018/19 Sentinel-2 satellite imagery was acquired over New Zealand and mosaicked into national snapshots of the country for each period. These snapshots were then compared to identify areas of forest loss which occurred during 2017 and 2018.

The objective of the deforestation mapping project was to field check these areas of forest loss by acquiring high-resolution aerial photos of each of the areas, classifying the current land cover in these areas, determining which areas have been replanted, and recording these findings in the spatial target layer provided by the Ministry. The project consisted of two primary activity areas:

1. acquire evidence images for the full extent of all forest loss targets

2. produce a spatial layer of forest loss target boundaries attributed with the evidence image name and replanting-related attribution (e.g. “fully replanted”; “partially replanted”; or “not replanted”) based on the evidence image.

To produce the final forest loss spatial layer, we also generated an intermediate layer which reports land cover at a 10m spatial resolution across each target. This target output resolution of $100m^2$ is referred to as super-pixels. This was the final output from a deep learning model trained on the imagery acquired in the project.

Our solution encompassed eight project phases run in a serial process to generate the evidence images and output data. The stages are depicted in Figure 1 and described at a high level below.

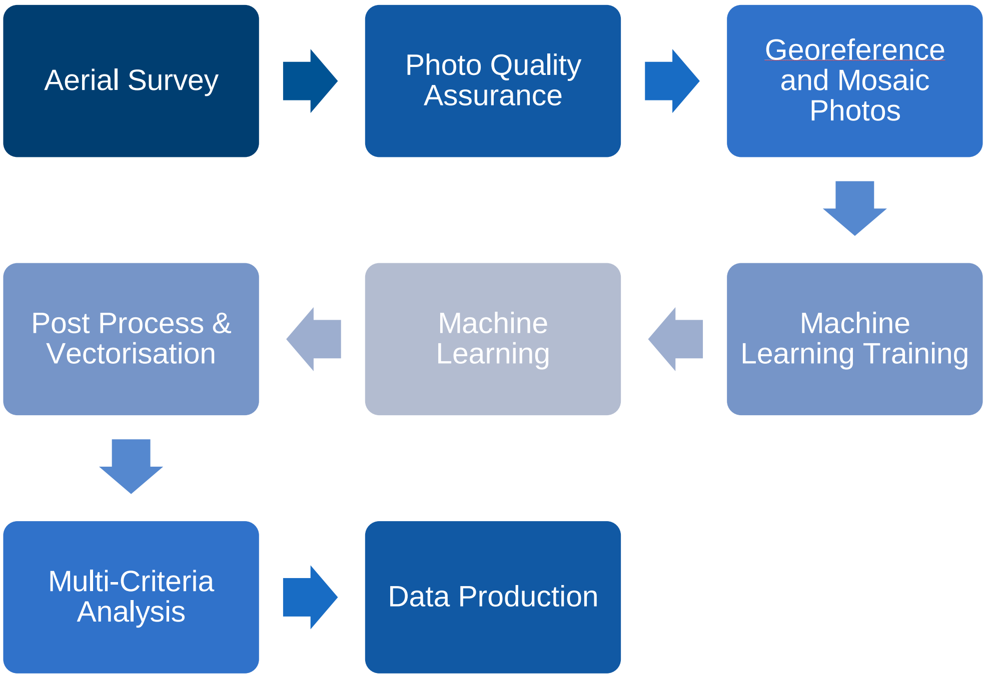


Figure 1 Solution workflow and processing pipeline

The aerial survey produced an individual photo mosaic for each deforestation target. Following image checks, these photo mosaics were then staged for input into a deep learning model. We used a patch-segmentation method to classify ground cover in patches of $100m^2$ within every target. This resolution was selected as it enabled us to review and determine land cover considering the typical spacing between trees in plantation forest. Our approach uses Active Learning or 'human in the loop' computing – a form of supervised machine learning that requires only the most informative samples for training.

From there we applied a filtering process to reduce noise from the model output. This resulted in a land cover map for each deforestation target inclusive of up to 10 classes including plantation seedlings, cutover, mature exotic, mature native, and pasture.

Finally, we applied a multi-criteria analysis to the sub-area polygons to assess the ratio of land cover within each target. The result of this was the production of a spatial layer of forest loss target boundaries attributed with the evidence image name and replanting-related attribution.

2 Aerial Photography

The aerial survey was managed by UAV Mapping NZ Ltd using Cessna 172 aircraft (operated by Action Aviation) fitted with a full frame digital, nadir camera. Flight planning was carried out using the Aviatrix flight management system (FMS) and we used live-tracking GPS to fly a predetermined flight path.

The camera was a Canon 5D SR full-frame CMOS sensor fitted with a 24mm lens. The aircraft height was 4,000 to 5,000 feet above ground level (AGL). As the aircraft approached the area to be captured the camera triggered automatically and captured the image(s).

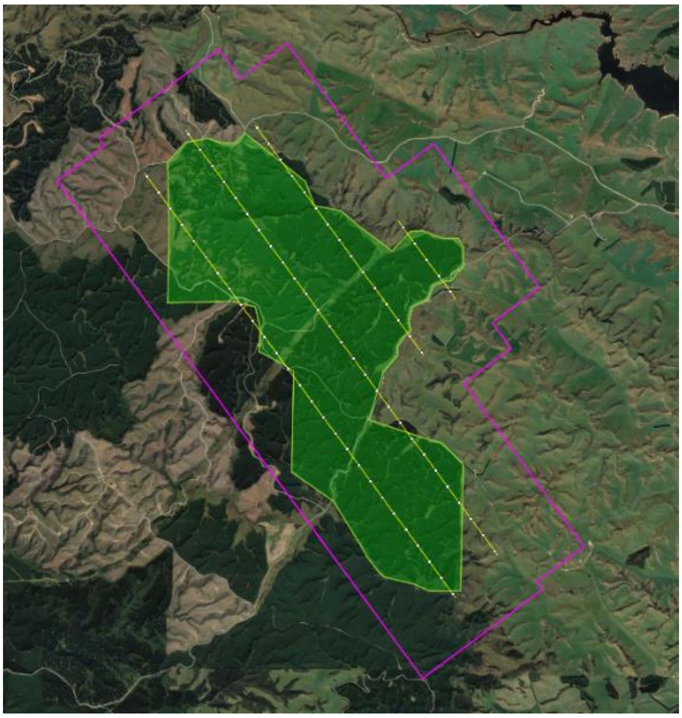


Figure 2 Flight plan over deforestation

A minimum of four nadir photos were captured over each target. Each photo was geotagged with WGS84 coordinates in decimal degrees, alongside detailed metadata including datetime, aircraft height, and pitch/roll/yaw.

The camera acquired 3-band nadir images (RGB) in Canon CR2 raw format to ensure maximum image data was available. These raw files were then colour-balanced to settings that would optimise the ML process downstream.

Most targets only required four photos while larger targets such as the example shown in Figure 2 required multiple flight lines and incorporated side 80% forward and 50% side overlap.

The aerial survey started on January 23, 2020 and completed on August 17, 2020. Due to the COVID19 lockdown the aerial survey was grounded between 25 March and 7 May (7 weeks). This resulted in some of the winter imagery including longer shadows than desirable. A summary of the aerial survey is provided in Table 1 below.

Metric	Quantity
Photos Captured (50mp 12bit raw images)	30,810
Mosaics Created	5,149
Acquired Targets	7469
Flying Days	55
Flight hours	226.2
Diesel consumed (litres)	6107

Table 1 Aerial Survey Summary Statistics

A further four targets were acquired using ground-based photography – three in Wellington and one in Auckland. A further 11 targets were not acquired due to weather and/or CAA restrictions.

Using the photograph overlap (mostly forward only) we generated a single photomosaic using an ortho-rectification process and these were then georeferenced using a semi-automated process within a GIS using existing orthophotography as a reference. The positional accuracy of the final 0.25cm GSD imagery is approximately +/- 20m.

Figure 3 Georeferencing process, new photo (left) and orthophoto reference (right)

All carbon emissions associated with the aircraft operations were offset on a forestry project on Banks Peninsula.

3 Machine Learning

Multi-scale convolutional neural network (CNN) models were used to classify each target boundary into land cover categories to a spatial resolution of $100m^2$. In deep learning, a convolutional neural network is a class of deep neural network that is most commonly applied to analysing visual imagery.

To train the models, we used Active Learning – a methodology used to achieve high accuracy models using only the most essential training inputs.

Shown in Figure 4, our Active Learning method employs a human-in-the-loop methodology to iteratively select and present images for labelling. The neural network models propose labels on the images that have the largest entropy (least confidently labelled or most uncertain).

In the training phase, these are reviewed and corrected by human experts and then added to the pool of labelled data for model retraining. This continues until all high entropy imagery is exhausted, which indicates that the model is now suitable for running against future, unlabelled data.

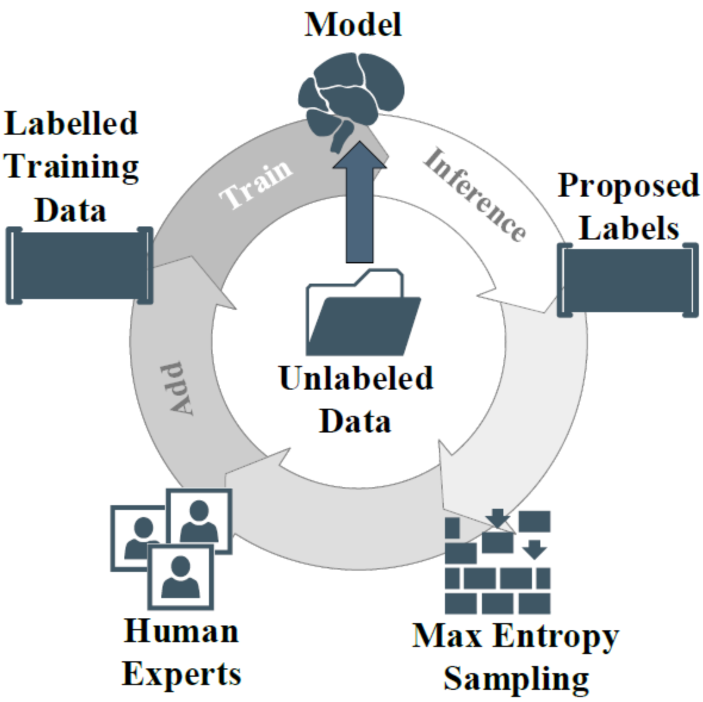


Figure 4 Active Learning Method

This max-entropy sampling method ensures only the most informative samples are reviewed and labelled by a human expert, leading to savings in human effort and processing time. This sampling method also mitigates the effects of sampling bias in the training set or from our reviewers.

Examples of the multi-scale training annotations used in the project are shown in Figure 5. In each example, the left-hand image shows an area of 70m x 70m with the red box representing the right-hand image, which is an area of 10m x 10m.

In total we captured approximately 15,000 multi-scale training annotations across all classes.

Plantation Seedlings *Mature Exotic*

Figure 5 Training Annotations

For the neural network models, a patch-segmentation method was trained on the multi-scale image chip pairs. Patch segmentation was used instead of U-Net (or similar) for semantic segmentation, because an image classification task allowed us to gather training annotations quickly and also land use classification makes more sense at this scale to reveal replanted seedlings separated by scrub and/or cutover.

We used Google's Inception V3-based neural networks which are pre-trained on the ImageNet dataset. These are well understood models and a good compromise between efficiency and performance. Transfer learning and fine-tuning was then applied to the neural networks, producing a fully connected classifier which combined the context view (70m x 70m) and a detailed view (10m x 10m) to classify 100m^2 image chips. Figure 6 shows the neural network design and the two scaled views used by the neural networks.

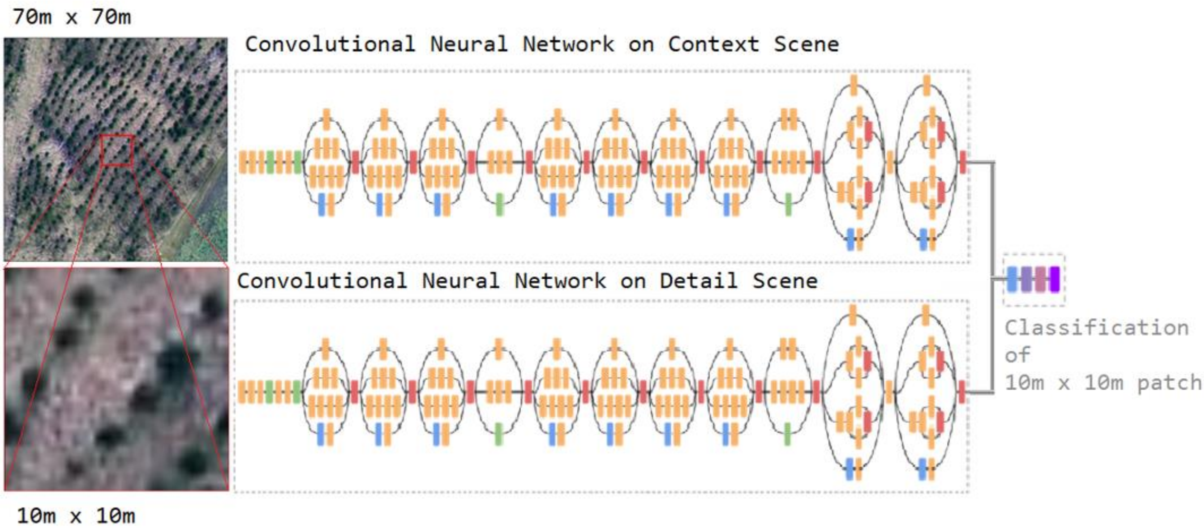


Figure 6 Neural Network Architecture

The context view is approx. 280x280 pixels, while the detail view is approx. 40x40 pixels. Both are resized to 299x299 for input into Inception. The diagram depicts the layers within the neural network where mathematical transformations are applied repeatedly against the previous layer output until a final prediction is made.

In the end the machine learning algorithm classifies all $100m^2$ patches within the target area image into one of the following classes.

1. Built Forest (cut site, access tracks)
2. Built Other (pavement, buildings)
3. Crop
4. Cutover
5. Exotic Regenerating Forest
6. Grass/Pasture
7. Horticulture
8. Mature Exotic Forest
9. Mature Native Forest
10. Natural Other (scree, riverbank)
11. Natural Regenerating Forest
12. Other
13. Plantation Seedlings
14. Shadow
15. Water

Table 2 Landcover categories

.

4 Model Accuracy

A confusion matrix is presented to examine accuracy of the machine learning model. The following table and discussion examine the confusion and accuracy of the primary class sub-area land cover machine learning model predictions relative to the validation data captured by our team. We have only included the major classes here, as there were insufficient examples of some of the sparser classes such as crop, exotic regeneration vegetation, natural damage, and water.

The validation data was not used to train the machine learning models, but was captured at the same time and using the same process as the training data. The data below is based on the raw output of the land cover machine learning model. As such, it will include sub-area pixels that are removed in subsequent GIS post processing.

In Table 3, the first row is the predicted class while the first column is the actual (validation) class at a super-pixel level. The number in each cell represents the number of super-pixels from the validation set classified by the model into output classes.

CLASS	Built Forest	Built Other	Cutover	Pasture	Mature Exotic	Mature Native	Natural Regen	Seedlings & Exotic Regen	Actual Count
Built Forest	14	7	11	3	0	0	0	1	36
Built Other	0	3	3	1	2	0	0	0	9
Cutover	3	1	259	16	0	6	9	3	297
Pasture	5	3	25	256	1	5	4	0	299
Mature Exotic	0	0	6	1	57	15	3	6	88
Mature Native	0	0	0	0	8	68	16	2	94
Natural Regen	0	0	36	1	2	5	27	9	80
Seedlings & Exotic Regen	0	0	14	2	0	6	0	69	91
Predicted Count	22	14	354	280	68	105	56	90	

Table 3 Confusion Matrix

The grey highlighted cell presents the number of super-pixels classified by the model into the correct class. For example, 259 of 297 Cutover super-pixels (87%) were correctly classified by the model into the Cutover class. We have also highlighted examples in yellow where the model has incorrectly labelled pixels. These are examples where the model finds it more difficult to distinguish between classes. The count of both actual (validation) and predicted super-pixels is included also.

Of note is the confusion between natural regenerating forest and cutover. This is to be expected, as there is a gradual transition from cutover to scrub, bush and regenerating forest and the division between these classes can be challenging for even a human reviewer to determine.

Overall, the combined model accuracy is **0.76 (76%).** This is calculated by dividing the number of correct predictions by the number of total predictions.

Given similarities in land cover visually and ambiguity in separating many of the gradational classes, this is a good result overall; which enables the model to be used as a detection system for change in combination with the multi-criteria analysis.

To explain the overall model performance and accuracy we have provided a summary of model classification statistics in Table 4.

CLASS	Samples	TP	FP	TN	FN	Precision	Recall	F1 Score
Built Forest	36	14	8	950	22	0.64	0.39	0.48
Built Other	9	3	11	974	6	0.21	0.33	0.26
Cutover	297	259	95	602	38	0.73	0.87	0.80
Pasture	299	256	24	671	43	0.91	0.86	0.88
Mature Exotic	88	57	13	893	31	0.81	0.65	0.72
Mature Native	94	68	37	863	26	0.65	0.72	0.68
Natural Regen	80	27	32	882	53	0.46	0.34	0.39
Seedlings & Exotic Regen	91	69	21	882	22	0.77	0.76	0.76

Table 4 Classification Statistics

This table presents several statistical parameters including:

- True Positive (TP) = the model correctly predicts the positive class
- False Positive (FP) = the model incorrectly predicts the positive class
- True Negative (TN) = the model correctly picks the negative class

- False Negative (FN) = the model incorrectly predicts the negative class
- Precision = TP / (TP + FP)
- Recall = TP / (TP + FN)
- F1 = 2 X ((Precision X Recall) / (Precision + Recall))Precision,

Recall and F1 scores provide additional measures of accuracy by class. Precision shows how precise/accurate the model is out of those predicted to be positive. Recall calculates how many of the actual positives the model capture through labeling it as True Positive.

The F1 score is a useful measure of overall model accuracy as it represents a weighted average of Precision and Recall.

In this model these values are lower for classes such as Built Forest and Built Other but relatively high for the more commonly occurring vegetation, pasture and Cutover classes. We expect a reasonably wide variation given the heterogeneity of these environments and the similar visual appearance of classes such as Cutover, Seedlings and Pasture.

The subsequent filtering and multi-criteria analysis within the GIS provided further generalisation value to deliver a refined sub-area land cover map used to determine final re-plant status.

5 GIS Post-processing

Following the machine learning phase, we applied a two-stage GIS process to eliminate scatter and noise from the inference results, clip, and vectorise and then apply a multi-criteria analysis to assign a dominant land cover and re-plant status. These two stages are detailed below.

5.1 Vectorisation

The deep learning models produced multi-class raster images at a 100 sq m super-pixel resolution. These images spanned the entire image footprint which extends beyond the deforestation target polygon. We applied a GIS post processing method to vectorise, filter out noise (speckle) and clip to the target boundary.

Once the raster was converted to a polygon feature class we applied an eliminate polygon process which eliminates a polygon by merging it with the polygon from the surrounding features that it shares the longest boundary with. In this case we removed polygons with an area size of less than 300 square metres. This removed the minor isolated land cover groupings which while interesting were not important to the final land cover determination at a target scale.

The final feature class was then clipped back to the target boundary. An example of this process is shown in Figure 7.

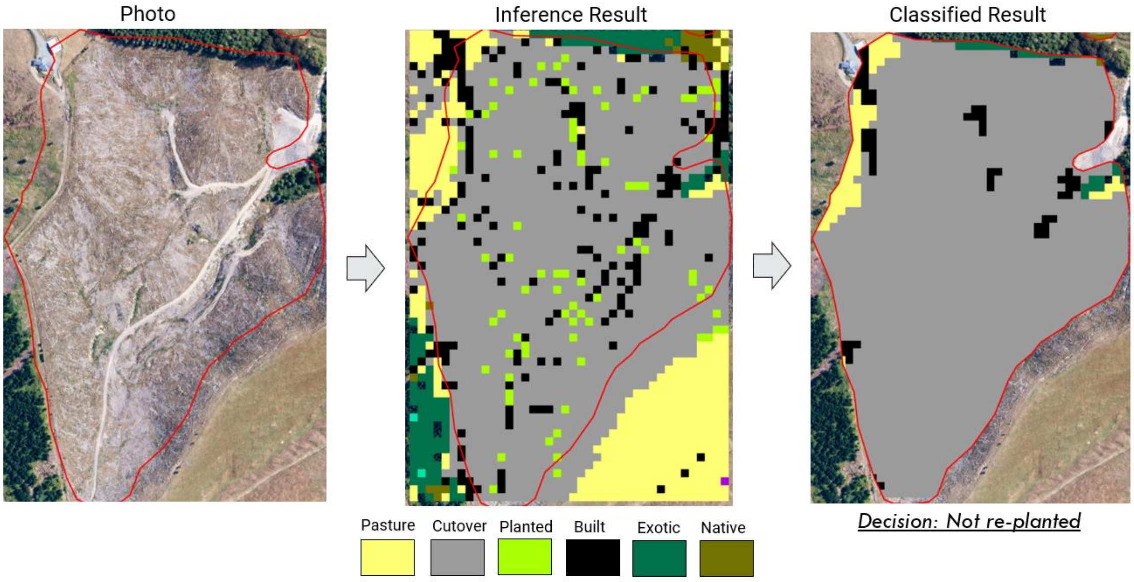


Figure 7 GIS Post-processing

The final polygon layer for each target holds the area of each class in square metres. This information is then used in the second stage analysis.

5.2 Multi-Criteria Analysis

Taking result polygons from the previous stage we applied a sequential decision process to classify the target polygon into dominant land cover categories broadly in line with the Emissions Trading Scheme - Geospatial Mapping Information Standard.

Each target polygon is tagged with a final classification based on the criteria and decision sequence shown in Table 5.

STEP	CRITERIA Exclude Shadow, Mature Native, Mature Exotic, Built Forest	DOMINANT LAND COVER
1	Built Other + Pasture + Crop + Horticulture > 80%	Fully deforested
2	Built Other + Pasture + Crop + Horticulture > 1 Ha	Partially deforested
3	Plantation Seedlings + Regen > 70%	Fully re-planted
4	Plantation Seedlings + Regen > 1 Ha OR > 30%	Partially re-planted
5	Cutover > 1 Ha or > 30%	Not re-planted
6	(Re)-include Mature Exotic > 1 Ha	Still Forest
7	Natural Damage + Natural Other > 1 Ha	Natural Adverse Event
8	All other	Unknown

Table 5 Multi-criteria decision process

We use a sequential decision process to make the final determination with polygons qualifying against the criteria being categorised into that dominant land cover. This is a sequential process designed to output attributes in order of importance to the review process. Deforested targets being of high interest followed by re-planted targets. Any targets not categorized are assigned to Unknown.

We note also that the four target sites photographed from the ground were assessed manually into a dominant land cover status. Several photos were classified manually due to extensive winter shadows.

6 Deliverables

The deliverables include the following:

A. an updated Target Layer (polygon feature class) where each and every target has been updated with a dominant land cover classification and replanting status based on the method described in this report

B. for every Target, the applicable evidence image(s) are provided and named to match the ORIGINAL_RECON_IMAGE_NAME attribute supplied in the Updated Target Layer (“Evidence Imagery”)

C. a file geodatabase containing:

   a. a polygon feature class of footprints and metadata for all aerial imagery supplied.
   b. a point feature class of locations and metadata for all terrestrial imagery supplied.

D. a land cover layer delineating Target sub-area and attributing each sub-area with its land cover extent

E. Python code including machine learning model and model parameters

F. a final report (this document)

7 Key Learnings

The following list of key technical findings and lessons learned are provided for future reference.

1. **Vertical Aerial Photography**

A single photograph with a photocenter GPS position from 5000 feet while capable of providing areal coverage of most targets did not enable satisfactory positional accuracy. Consequently, we captured multiple photos (at least four in most cases) using forward overlap (and side overlap in the case of larger targets). The wider swath width enabled more context of the targets to be understood.

2. Geo-referencing

An ortho-rectification process was used to produce an image mosaic which was then repositioned manually where required within a GIS using existing orthophotography as a reference. In the case of larger targets, the GPS data was accurate enough to enable automated placement. Accurate flight planning is crucial to balance and optimise the total image footprint against flight time and post processed GPS data. Overall, we estimate the average positional accuracy of the imagery to be +/- 20m.

3. **COVID19**

The pandemic required the survey team to validate all health and safety systems and documentation and ensure alignment with the latest legislation. The team took on additional protective personal equipment, sanitizer and cleaning regime as operations commenced after the national lockdown. A consequence of the delay were additional shadows in the imagery captured between May and August.

4. **Machine Learning**

Patch segmentation using multi-scale CNN models proved to be a successful method to use in combination with 0.25m GSD RGB (optical bands) vertical photography. We used Google's Inception V3-based neural networks which provide a good compromise between efficiency and performance. Over 15,000 training annotations were acquired regionally and seasonally. Training needed to be continually captured throughout the project to build a model suitable for the entire country and temporal range of the imagery. These were all acquired using the multi-scale views within an active learning interface. In some sparsely occurring classes such as crop, horticulture and natural damage there were insufficient examples to deliver a high-confidence classification. In future using terrain from outside the target area might be a way to overcome this.

5. Land cover Classes

We started with 15 training classes including categories such as “Cutover-Cleared” and Cutover-Scrub” and finished with 13 classes excluding shadow. In the end a more generalized definition of Cutover delivered better results. Winter shadows also presented a problem to the models and rather than including it, we elected to assign it to NODATA. In future, should winter imagery be required, this would be better handled using a multi-season environmental imagery correction method to equalize brightness, contrast and colour.

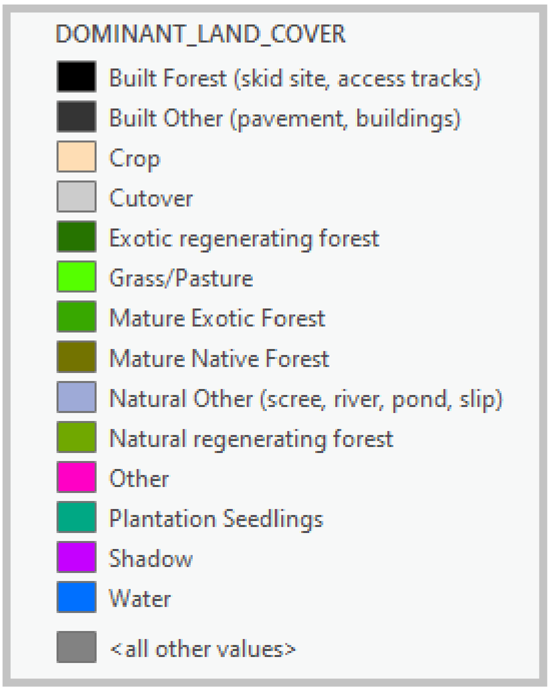
6. GIS Post-processing

The two-stage post-processing approach worked well. We elected to filter out small, isolated polygons of three or less super-pixels using an eliminate process which eliminates a polygon by merging it with the polygon from the surrounding features that it shares the longest boundary with. The final multi-criteria analysis was also developed iteratively over the course of the project with the final decision sequence designed to alert the Ministry to those targets that may contain land use change of at least 1 hectare.

Appendix 1

Example Classification Results

Examples of the sub-area classification are provided below. Note the inference result is generated at a 10m x 10m spatial resolution.



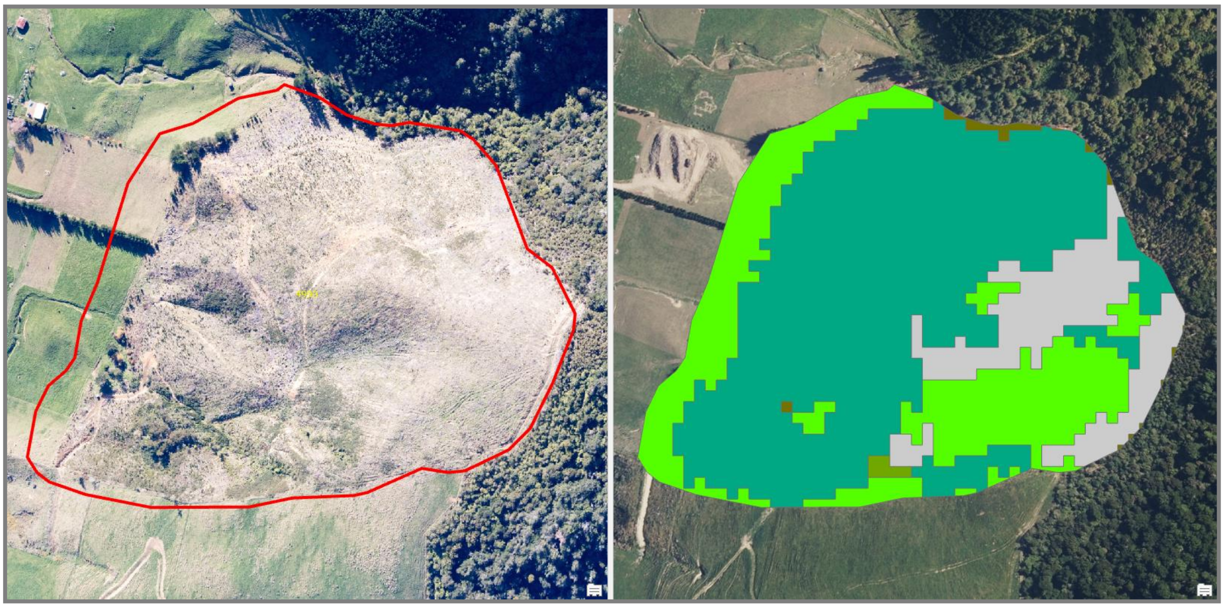
In each case the aerial photograph is shown at left and the classified polygon layer at right.

The polygon layer is the final delivered sub-area data set which has undergone GIS post-processing.

The classification legend is shown here. For each target we have additionally provided a summary of the following:

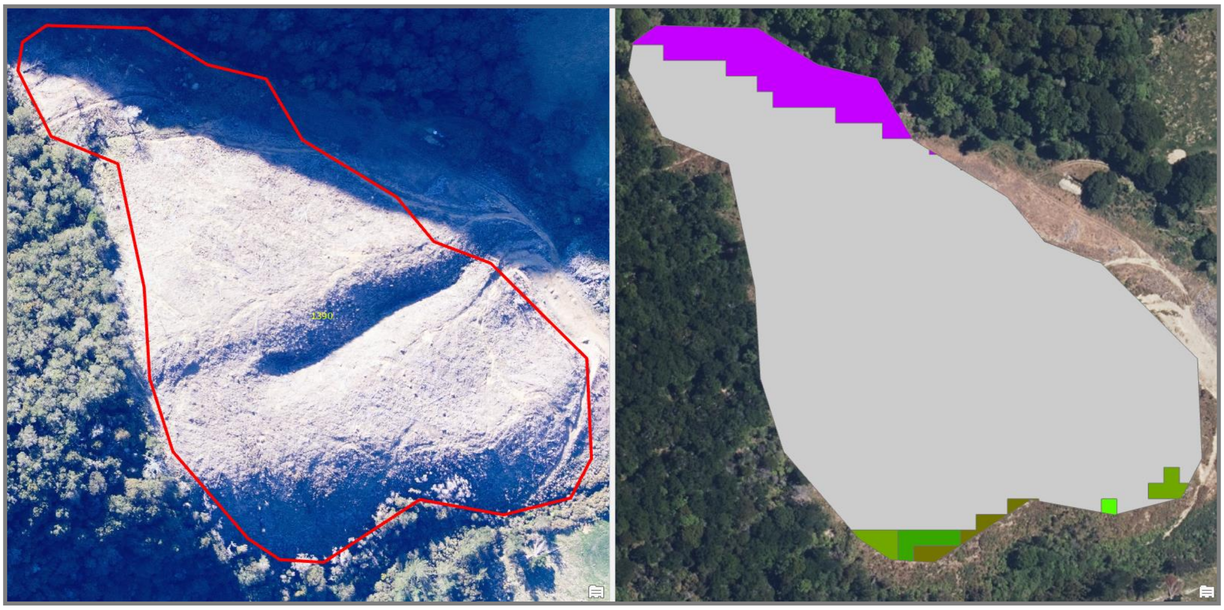
- target area in hectares
- destock year
- dominant land cover
- replant status

LKR_2018_MAR_0104 – MARLBOROUGH DISTRICT



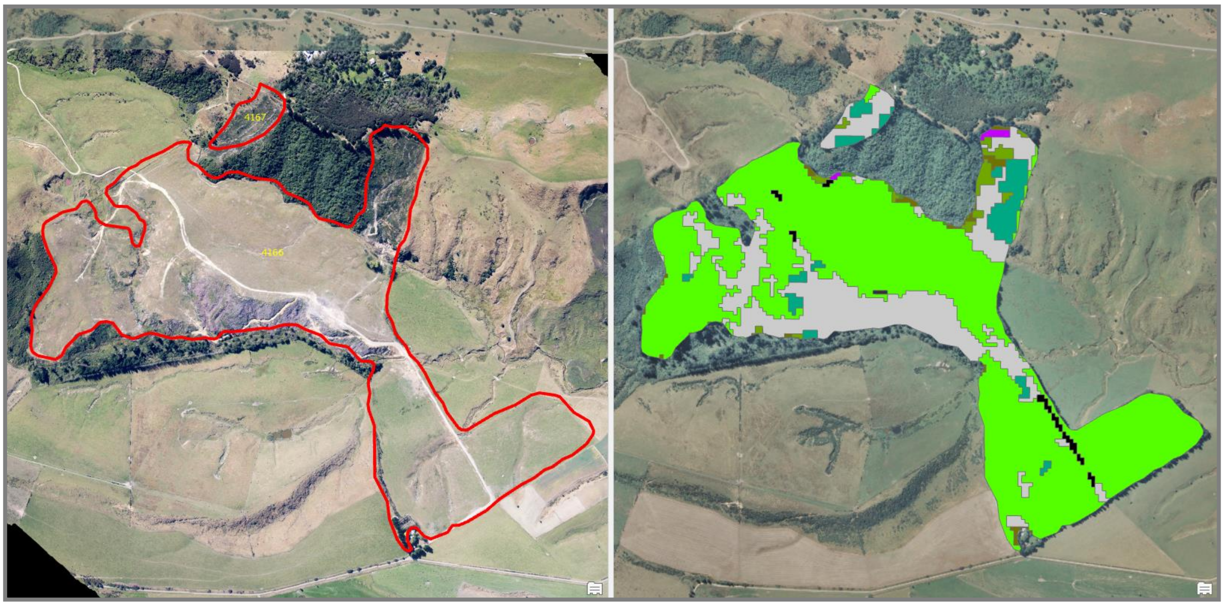
Area (ha)	Destock Year	Dominant Land Cover	Replant Status
14.1	2017	Plantation Seedlings	Partially Deforested

IAP_2014_MAR_0325 – MARLBOROUGH DISTRICT



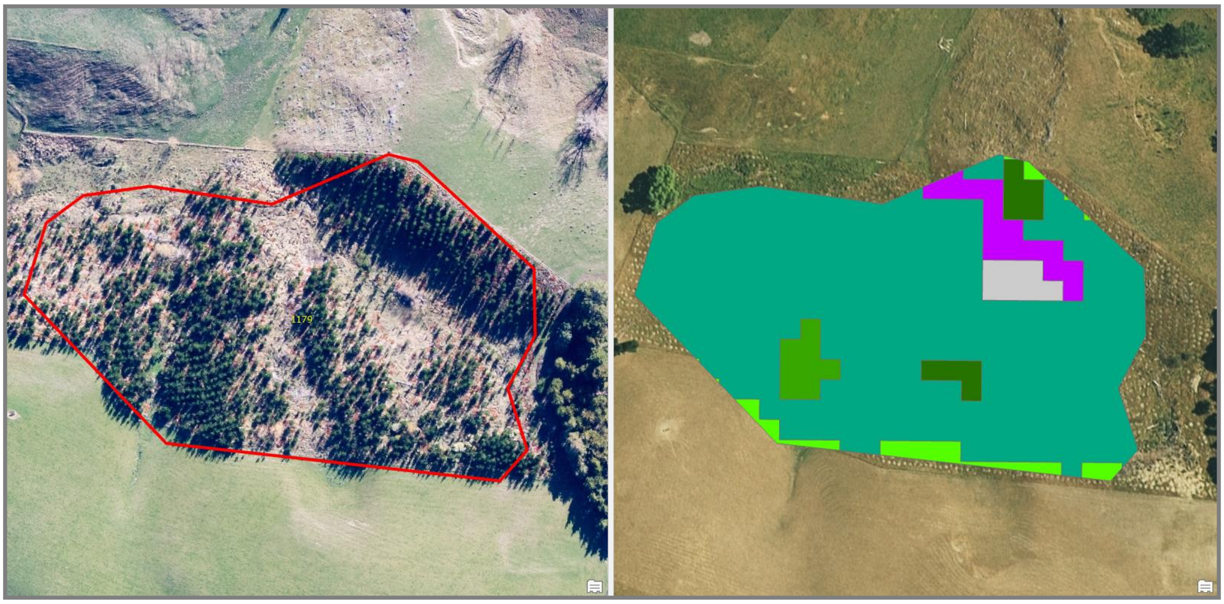
Area (ha)	Destock Year	Dominant Land Cover	Replant Status
6.6	2014	Cutover	Not re-planted

LKR_2018_CAN_0171 - CANTERBURY



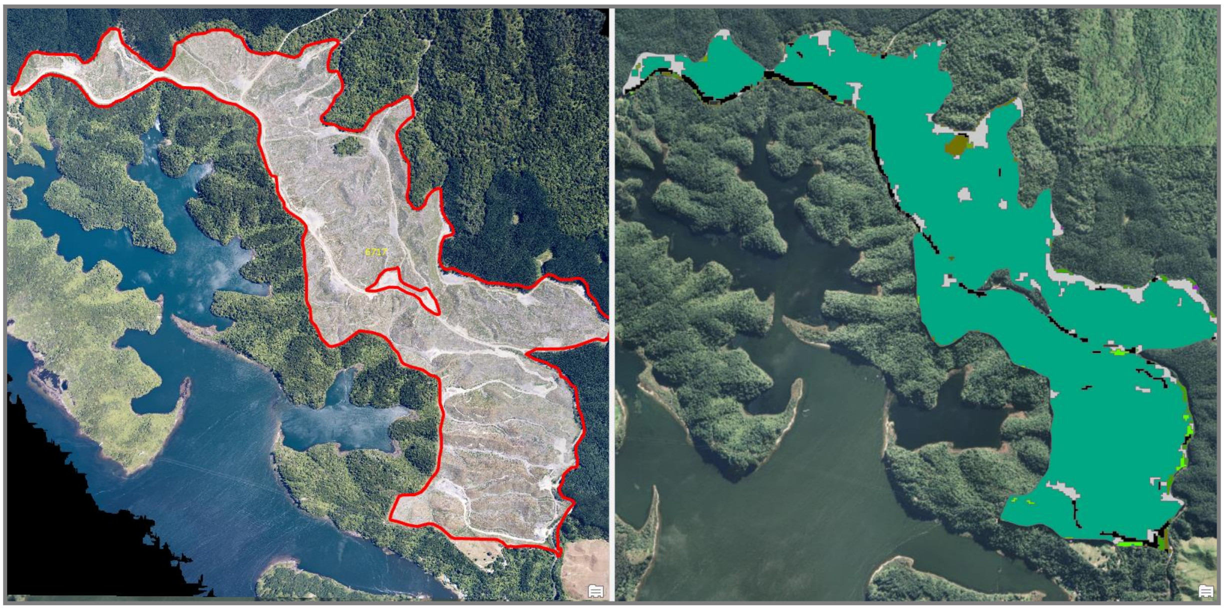
Area (ha)	Destock Year	Dominant Land Cover	Replant Status
58.5	2017	Grass/Pasture	Partially deforested

IAP_2014_HWK_0313 – HAWKES BAY



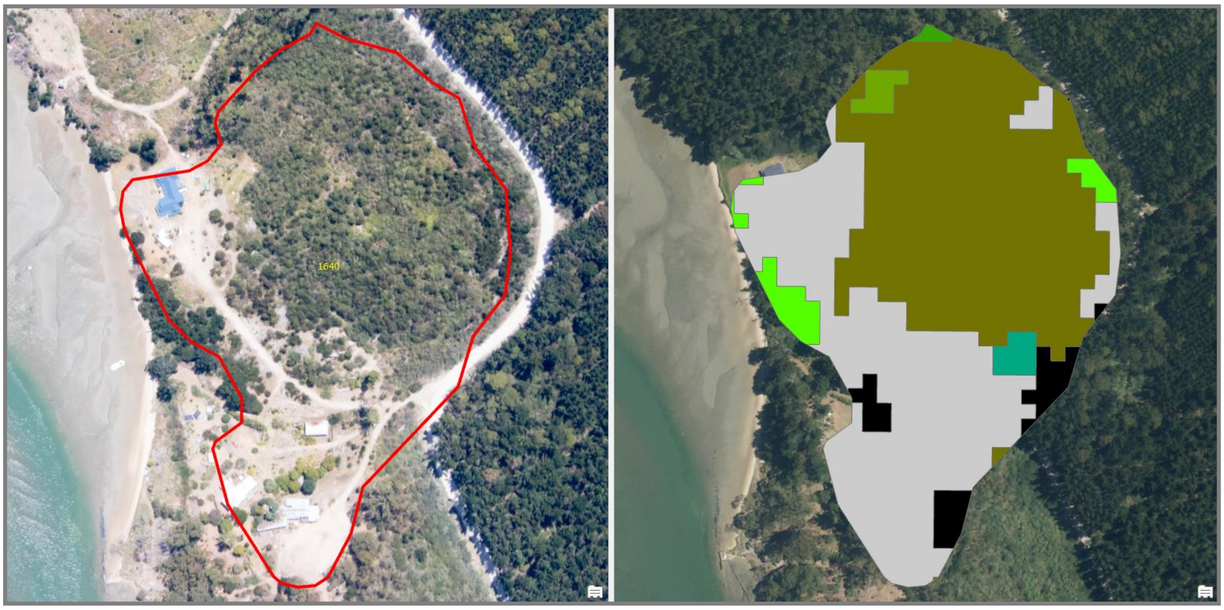
Area (ha)	Destock Year	Dominant Land Cover	Replant Status
2.9	2014	Plantation seedlings	Fully re-planted

LKR_2018_BOP_0155 – BAY OF PLENTY



Area (ha)	Destock Year	Dominant Land Cover	Replant Status
178.8	2017	Plantation seedlings	Fully re-planted

IAP_2016_BOP_0440 – BAY OF PLENTY



Area (ha)	Destock Year	Dominant Land Cover	Replant Status
6.1	2015	Mature Native	Not re-planted

Appendix 2

Machine Learning Code

1. The project code was copied to the Ministry's GutHub repository.
2. The repository consists of
    a. Python code and a neural network model weights file.
3. The python script will run the ML land classification inference over RGB aerial imagery to produce a raw land cover classified raster.
4. The models are trained on vertical aerial photography with a spatial resolution of 0.25m.
5. The repository contains a requirements.txt file that lists the python libraries required to run the code and a README.md file that describes how to run the ML inference.

Lynker Analytics Ltd
Suite 6, Level 2
10 Hutt Road
Petone
Lower Hutt 5012
New Zealand

Carbon Forest Services
60 Cuba St
Wellington 6011
New Zealand

UAV Mapping NZ Ltd
14 Smiths Rd
RD-1
Ohakune 4691
New Zealand