AVERAGE AMOUNT OF INFRASTRUCTURE REQUIRED TO HARVEST SMALL-SCALE WOODLOTS IN NEW ZEALAND AND THE INFLUENCING FACTORS.

Final Report

Jacob Allum - 07/10/2020
University of Canterbury

Table of Contents

Acknowledgements:

I would like to thank all of those who have assisted me throughout the scope of this project. In particular, I would like to thank Rien Visser and Stephan Hoffman for coordinating the ENFO410 forest engineering research course and assisting us all with our projects along the way. Special thanks also need to go to Campbell Harvey and Glen Murphy, who both helped me in the development stages of my project, as well as answering any questions I had at various times. Thank you to Vega Xu, Bruce Manley, and the University of Canterbury for supplying the ArcGIS shapefiles of mapped woodlots throughout New Zealand, that I used throughout this study. Finally, I would like to thank the family of John Dey and the NZIF Foundation for their generous financial support that enabled me to complete this project. This project would not have been possible if I had not received the quality guidance and assistance from those mentioned above.

Abstract:

Small scale woodlots between 0 and 500 hectares are increasingly accounting for a higher proportion of the total forested area of New Zealand. Currently, there is no record of national averages for the amount of infrastructure required to harvest a small-scale woodlot in New Zealand. That being; the road density and landing size required on average to complete a clear-cut harvest, along with the average area that each of these landings' services.

This study quantified national averages around these infrastructure values and evaluated what factors influence these averages. Using ArcGIS and Google Earth, roads and landings were mapped throughout a sample of 96 woodlots across the 9 wood supply regions within New Zealand. These programs were also used to assist in finding the area, average slope, length/width ratio, boundary complexity, extraction method, and soil type for each of these 96 samples. Average values were then calculated for the infrastructure variables, and through multiple multivariate regression analyses and statistical analyses, the other variables listed were assessed for their influence on infrastructure.

The infrastructure averages gained were as follows; road density (0 m/ha samples included) = 25.2 metres/hectare, road density (0 m/ha samples excluded) = 29.9 metres/hectare, landing size = 3000.1 m^2, and landing service area = 12.8 hectares/landing. The first road density value included 15 woodlot samples that had no internal roading; the second is the average road density with these 0 metres/hectare samples removed. From strongest to weakest, the following relationships were found. Road density had a significant relationship with; length/width ratio, average slope (%), and boundary complexity. Landing size had a significant relationship with; average slope (%), length/width ratio, and area (ha). Finally, landing service area had a significant relationship with; boundary complexity and area (ha).

The infrastructure results from this study may have several uses to industry including; assisting in initial harvest planning, helping with cost reduction, assisting in due diligence assessment, through being used as inputs to models and cost analyses, and by supplementing further research.

1. Introduction

This study will aim to quantify average values for the amount of infrastructure required to clear-cut harvest small-scale woodlots up to 500 hectares in New Zealand. Along with this, it will evaluate the driving factors behind these averages through the use of statistical analyses and multiple multivariate regression analyses. The infrastructure averages gained will be for; road density, landing size, and landing service area. The factors assessed for their influence will be woodlot area (ha), average slope (%), length/width ratio, boundary complexity, extraction method, and soil type. Through not only having applicable values but knowing what influences them, it could greatly benefit harvest planners, cost analysts or due diligence assessors. For example, if there is clear evidence on a factor that decreases road density, this could be manipulated to in turn decrease costs. With the large quantity of wood stored in small-scale woodlots approaching optimum harvest age due to increased planting in the 1990s, this study could greatly benefit a large proportion of the New Zealand forestry industry.

2. Literature Review

2.1 Small-scale woodlots and the importance of information.

Small-scale woodlots are a vital component of New Zealand's forestry industry. The size definition of small-scale woodlots varies between literature. Two studies quantifying the area of small-scale forest estates in differing wood supply regions of New Zealand define a small-scale woodlot estate as less than 1000 hectares (Manley, Morgenroth, Xu, & BForSc Students, 2020; Manley, Morgenroth, Visser, & BForSc Students, 2017). The Wood Availability Forecasts published by the Ministry for Primary Industries also define small-scale woodlots as less than 1000 hectares (Indufor Asia Pacific, 2016). Forest Growers Research Limited funded the production of a document outlining how to market and harvest small-scale woodlots for profit within a New Zealand environment. It stated that a woodlot should be defined as less than 500 hectares (Visser & Murphy, 2019).

As of the 1st of April 2019, small-scale woodlots between 0 and 500 hectares account for 27.3 % of the total plantation area of New Zealand; this equates to 462,621 hectares (Ministry for Primary Industries (MPI), 2019). According to the Ministry for Primary Industries, the primary wood supply regions in New Zealand are as follows; Northland, Central North Island, East Coast, Hawkes Bay, Southern North Island, Nelson/Marlborough, West Coast, Canterbury, and Otago/Southland (Ministry for Primary Industries (MPI), 2019). The percentages of the 462,621 hectares of forest by wood supply regions are as follows (Ministry for Primary Industries (MPI), 2019).

Table 1: Area and percentage of total small-scale woodlots by wood supply region.

Wood Supply Region:	**Area of small-scale woodlots 0 – 500 (ha):**	**% of total small-scale woodlots (%):**
Northland	55,391	12.0%
Central North Island	83,194	18.0%
East Coast	28,620	6.2%
Hawkes Bay	33,135	7.2%
Southern North Island	88,570	19.2%
Nelson/Marlborough	53,175	11.5%
West Coast	3,595	0.8%
Canterbury	49,971	10.8%
Otago/Southland	66,972	14.5%

Since 2014, the fourth year Bachelor of Forestry Science students at the University of Canterbury have plotted all the woodlots up to a size of 1000 hectares within 7 of the nine wood supply regions, with the Central North Island and Northland to be completed around the beginning of August 2020. Reports that outline this procedure were published in 2017 and 2020 ((Manley, Morgenroth, Visser, & BForSc Students, 2017); (Manley, Morgenroth, Xu, & BForSc Students, 2020)). Aerial imagery was used to determine the location of woodlots within each wood supply region. For all regions apart from Canterbury, Land Information New Zealand (LINZ) aerial imagery was used. This aerial imagery ranged from resolutions of 0.125 metres to 0.75 metres. In the Canterbury region, aerial photographs from the environment Canterbury web map tile service were used, these vary in resolution from 0.4 to 0.75 metres. In all cases, the woodlots were then cross-referenced with Google Earth to check the status of the stands. To meet requirements for mapping the forested area had to be over 1 hectare, and greater than 30 metres wide. These rules were only relaxed when there were contiguous small blocks that added to over 1 hectare. All mapping was done to a scale of 1:4,000 or greater. Once the BForSc students completed mapping, quality control was completed on the results by experienced postgraduate students. All line-work was verified, and checks were made to ensure all small-scale plantations had been included. This process ensured accuracy and minimised errors due to omission and commission. The University of Canterbury has collated the results of these studies into a GIS shapefile that can be overlaid on aerial imagery to determine woodlot location along with area and perimeter.

Over recent years information and factors that relate to the planning and costs of harvesting a small-scale woodlot have become increasingly important. The infrastructure required for harvesting is one of these factors. Because much of New Zealand's plantations are on steep, erosion-prone, challenging terrain with a considerable distance to nearest ports, infrastructure construction and maintenance are one of the most critical drivers in the overall economics of a harvest operation. Along with this, planning is a crucial component to ensuring an infrastructure programme is successful (New Zealand Forest Owners Association Incorporated, 2012). The increased importance of factors relating to these costs and planning is due to many elements. Firstly, the increase in the quantity of wood being produced from small woodlots. As of March 2019, small forests accounted for 40% of the total harvest (Forest Owners Association (FOA), 2019). According to wood availability forecasts produced in 2016, from 2020 the potential wood available from small-scale forest owners will increase to around 15 million m^3 per annum through to 2035. This is largely due to small-scale forest growers who established forests during the planting boom in the 1990s (Indufor Asia Pacific, 2016). This land planted in the 1990s maturing has also meant that the portion of small-scale woodlots with respect to the entire forested area in NZ is becoming important for wood production (Manley, Morgenroth, Xu, & BForSc Students, 2020). On top of this, if forestry remains attractive for carbon forestry and to farm owners looking to diversify their farm income the importance of woodlots to the national supply will likely increase even further (Visser & Murphy, 2019). The global climate, with regards to COVID-19, has also had flow-on effects. A monthly market report for April 2020 discussed that in April 2019, China's daily usage rate was over 90,000 m^3, as of April 2020, it is just 50,000 m^3 (Laurie, 2020). With China accounting for a very large proportion of New Zealand's log exports, and COVID-19 being expected to disrupt exports for the next two quarters (Ministry of Primary Industries (MPI), 2020) this will likely have a significant impact on log price. Average export log price overall log grades has fallen from 166 $NZ/JAS m^3 FOB to 134 $NZ/JAS m^3 FOB between June 2019 and March 2020 (Ministry for Primary Industries (MPI), March, 2020). Small-scale woodlots are often more expensive to harvest than large forest estates due to them not already having the existing roading and infrastructure in place that large forest estates do (Indufor Asia Pacific, 2016). The effects of this combined with low log prices and poor market conditions will make it significantly more important for costs to be kept low during the harvesting of small woodlots and for accurate cost analyses to be completed to evaluate the profitability of harvesting at this time, as opposed to waiting till market conditions improve. With infrastructure being one of the highest costs of harvesting a woodlot, averages behind the amount required could greatly assist with planning and costing relating to harvesting.

2.2.0 Infrastructure required to harvest woodlots.

Due to the 1990's planting boom, a large proportion of wood supply within New Zealand is in its first rotation. This means that a significant amount of infrastructure will need to be developed prior to harvesting. The aim of forestry infrastructure is ultimately about providing suitable access at the appropriate service level for forestry operations (New Zealand Forest Owners Association Incorporated, 2012). A study similar in nature to this proposed study was completed in 2014 by a

University of Canterbury student who defined infrastructure as “unproductive area” and aimed to assess the level of unproductive area in harvest sites within the South Island of New Zealand (Petherick, 2014). Unproductive area was found as the area of roads and area of compacted landings. Landings were defined as; “compacted area due to earthworks that is obviously being used as a landing”. The study concluded that from 41 sampled sites around the South Island, there was an unproductive area ratio of 4.82%, which means that 4.82% of the total forest area was necessary infrastructure for harvesting the woodlot. The required number of samples to achieve various confidence levels were calculated using Cochran’s formula. It also made use of weighted averages to determine the proportion of samples that will be undertaken within each region. The study identified total harvest area as a driver for unproductive area but did not manage to determine if harvest method influenced it. Downfalls of this study were; the limited success of the regression analysis that failed to identify any main drivers of this unproductive area other than harvest area, the unproductive area classification did not allow the evaluation of just landings or roads which both contribute to very different costs and planning requirements, and it was also a requirement that the estates evaluated had a minimum of 8 landings as this was thought to ensure a minimum size of 100 hectares. This is a downfall of the study as over 85% of woodlots less than 500 hectares in size are less than 40 hectares in area (Visser & Murphy, 2019), showing they should be included. Within the bounds of this study it mentioned a similar study was completed on the North Island by A. Mabbazza in 2014 and found a 4.2% unproductive area, this reinforces the accuracy of the 4.8% result as the values are fairly similar. Unfortunately, the North Island study was unpublished and therefore, unable to be located to verify the results.

2.2.1 Landings:

If landings are not located, designed or managed adequately, it can result in major consequences for safety, environment, production, quality and value recovery (New Zealand Forest Owners Association Incorporated, 2012). With regards to the infrastructure required for harvesting a woodlot, landings are one of two main components. Landing construction costs range from $4000 to $7000, depending on the size and design (Visser, Spinelli, & Magagnotti, 2010). Common landing layouts are; drive-through landings, roadside landings, spur road end landings, and split-level landings (New Zealand Forest Owners Association Incorporated, 2012).

One of the most cited studies conducted on landing sizes in New Zealand measured 142 landings with GPS in 2010 (Visser, Spinelli, & Magagnotti, 2010). Twelve were recently constructed, 38 were live, and 92 were older. The study found the average landing size within New Zealand operations to be 3,900 m^2 and samples ranged between 1,370 to 12,450 m^2. A landing was defined as any area that had been built and involved the removal of topsoil, was compact, flat and contiguous. If a road went through a landing, it was included in the landing area; if it ran beside the landing, it was excluded. Through asking forest managers to provide information on the respective variables, the study evaluated the effect of; landing age, type of operation (ground-based or hauler), type of processing (manual or mechanical), type of log loader used (front-end or knuckle-boom), number of log sorts, and daily productivity. Average slope was also evaluated in ArcGIS using 50 – 100 m radii circles from the centre point of the landing. It found that the two main factors influencing landing size are the number of log sorts and production level. Also, it found that landings serviced by front end loaders were on average 1,100 m^2 larger than landings serviced by a grapple loader. An interesting finding from the study was that used landings are on average much larger than recently constructed unused landings, this is reinforced in the NZ Forest Road Engineering Manual which states that a used landing is on average 900 m^2 larger (New Zealand Forest Owners Association Incorporated, 2012). This means that the company design cannot be used in this study to evaluate landing size as the design size differs to the size of landings actually used to harvest a woodlot.

The findings from the 2010 study can be compared to another study from 1987 that also evaluated the specific factors influencing landing size within New Zealand (Raymond, 1987). Raymond measured 50 landings across four different regions: Auckland/Coromandel, Bay of Plenty, Nelson, and Otago/Southland. The average landing size was found to be 2900 m^2, 1000 m^2 less than the average landing size from the 2010 study. The study also concluded productivity and loader type influenced landing size. It also found that soil type was a factor contributing to landing size. Significant differences were found between clay/loam types (typical of North Auckland and Otago/Southland) and pumice/sand types. Pumice/sand soils are free-draining, stable and easily workable, in turn tending towards bigger landing sizes. Clay type soils are more difficult to work and consolidate; this increases costs and creates pressure to decrease landing size. Interestingly, although hauler and ground-based methods have different requirements regarding space on landings, there was no significant difference found between the landing size of ground-based operations compared to hauler operations.

The New Zealand Forest Road Engineering Manual also states that landings may be influenced by soil type, types of machinery used, the number of log sorts, machinery production, and other factors (New Zealand Forest Owners Association Incorporated, 2012). This further supports the findings around the influences of landing size in the 1987 and 2010 studies.

A limitation of the above studies when comparing them to this study is that they evaluate landings within a range of forest sizes, including large forest estates. It will be useful to compare these average landing size values from the 1987 and 2010 studies to the value gained from this study to see if the exclusion of large-scale estates influences the average value.

2.1.2 Forest Roads and Density:

There are three main types of forest roads; these are Spur roads, which are short-term, low standard roads and generally carry less than 20 heavy vehicles per day (hvpd). Secondary roads, which are unsealed, permanent roads constructed to a high standard, they typically serve multiple operations and carry between 20 and 80 hvpd. Arterial roads are hub roads in major forests and are likely to carry truck traffic of more than 80 hvpd (Brown & Visser, 2018; New Zealand Forest Owners Association Incorporated, 2012).

Road density is the number of linear metres of road per hectare. For a given forest there will be a road density that minimises costs relating to construction, maintenance and timber extraction, this is the optimum road density (Ryan, Phillips, Ramsay, & Dempsey, 2004). There is a significant amount of literature published on optimum road densities; however, in terms of average road densities, the literature is limited.

A lot of the optimum road density studies are based upon costs, such as two studies completed in different forests in Iran. Both studies calculated skidding and roading costs and used these to find the optimum road density and then compared different skidders. Both studies used a similar skidder, the TimberJack 450C. The study completing the analysis in the Dalak Kheyl forest-Hycranian zone found the optimum road density to be 2-4 m/ha for a TimberJack (Rafiei, Lotfalian, Hosseini, & Parsakhoo, 2009), whereas the study completing the analysis in Kheiroudkenar forest found the optimum density using the TimberJack skidder to be 8.8 m/ha (using a one-way skid system) and 5.8 m/ha (using a two-way skid system) (Ghaffarian & Sobhani, 2008). These two studies are within the same country and using the same skidder; this shows that there is considerable variability between optimum road density within a forest setting as costs differ dramatically between forests. This shows that in terms of applicability in a range of scenarios, optimum road density for a specific forest is very limited.

A 2018 study by Brown & Visser used a targeted survey of active forest roading managers. It aimed to help better understand the characteristics of the forest industries current road construction programme. Major problems that were identified in managing their design programmes were; planning, designing and constructing infrastructure well in advanced of harvesting and controlling construction costs in steep terrain (Brown & Visser, 2018). A national average for road density could greatly help in the planning stage. The study also found that of new road construction, 63% will be spur roads, 34% will be secondary roads, and 4% will be built as arterial roads. The different types of roads have different construction costs. The forest managers were asked to estimate the average cost to build roads. The results found that spur roads on average cost $72,000/km, and secondary roads on average cost $90,000/km. However, the results from this study are limited, as only 3 of the respondents were from woodlots as opposed to large scale forest estates. Of the woodlot roads, lower-cost spur roads were found to cost $40,000/km or less. This reflects the lower road design standards for lower traffic volume, characteristic of woodlots. Another interesting finding of the study was that most road construction companies can construct a road off of a line plotted on a map.

Roading costs depend on length, standard, terrain and soil type, as well as access to the aggregate (Visser & Murphy, 2019; Brown & Visser, 2018). Also, depending on the composition of soil and rock in the earthworks, the ease of construction, cost of the job and environmental effects can be greatly affected (New Zealand Forest Owners Association Incorporated, 2012).

According to a forest road manual published in 2004 by COFORD, the National Council for Forest Research and Development in Ireland, Road density is related to the planned harvesting and extraction methods. Along with this, the overall cost of extraction within a forest setting is influenced most by road density and type of extraction machinery that will be used. (Ryan, Phillips, Ramsay, & Dempsey, 2004).

2.2 Aerial Imagery

The New Zealand Forest Road Engineering Manual lists freely available geospatial information and aerial imagery that can be used for the purpose of planning infrastructure.

- From Land Information New Zealand (LINZ) (www.linz.govt.nz); Map sheets, along with their underlying GIS datasets for the New Zealand Topo50 and Topo250 series maps. Historical orthophoto images can also be found (New Zealand Forest Owners Association Incorporated, 2012).
- From Google Earth (www.earth.google.com) or World Wind (www.worldwind.arc.nasa.gov); Satellite and aerial photographs for the whole of New Zealand. However, the terrain data is limited by the underlying digital terrain model, and the image quality will vary. (New Zealand Forest Owners Association Incorporated, 2012).
- From programs such as ECan GIS (www.ecan.govt.nz); Online GIS. These online applications often show additional features such as legal boundaries, utilities and natural resources.

2.3 Accuracy of Google Earth:

Google Earth is the most used internet service that provides a global collection of georeferenced satellite imagery. It allows humans to determine between major natural land cover classes (Farah & Algarni, 2014). Measuring tools within Google Earth Pro can be very useful for measuring features of aerial imagery displayed. The tools make measuring a line, path, polygon, circle, 3D path and 3D polygon very easy. There is currently no single measure of image or measurement accuracy listed by Google Earth (Harrington, et al., 2017). In order to validate using the measuring tools in scientific analyses, different studies have been completed on the accuracy of the measurements obtained.

One study aimed to quantify and compare the accuracy of using conventional methods such as measuring wheels and tape measures, with the tools within Google Earth. The purpose of this was to evaluate the effectiveness of using Google Earth Pro in accident reconstruction (Harrington, et al., 2017). From 68 locations within 25 states and provinces in the USA, Canada and Australia, 1305 unique measurements were compared. Current and historical satellite images were used for the Google Earth Pro measurements. For off-road measurements, the study found an average error of 1.61%, for on-road measurements, an average error of 1.41%, and for curved path measurements an error of 1.73%. It was found that as the length of the measurement increased, error rate generally decreased. Error rates were also found to be consistent between historical and most recent imagery.

Another study completed in 2014 evaluated the positional accuracy within Ridyah, Saudi Arabia (Farah & Algarni, 2014). This study compared the accuracy of Google Earth to a GPS unit. Nine control stations were set up, and readings using a LEICA-SR530 receiver were taken at each station for 30 minutes, using a 10-second interval. When compared to Google Earth, it concluded that the root mean square error (RMSE) of Google Earth imagery is 2.18 metres and 1.51 metres for the horizontal and height coordinates, respectively. A similar 2013 study completed in Khartoum State, Sudan, compared measured coordinates of points using Google Earth with a Trimble 1800 surveying GPS, over 16 checkpoints. The horizontal accuracy was determined to be 1.80 m, and vertical accuracy was determined to be 1.73 m, fairly similar to that of the 2014 Ridyah study. Overall, both studies confirmed that Google Earth is a powerful tool for investigation and studies to suitable accuracy.

3 Objectives:

The primary objective of this study is to quantify values for the average amount of infrastructure required to harvest small-scale woodlots between 0 and 500 hectares. That being; the average road density and average landing size, along with the average area each of these landings' services. The second objective is to evaluate what factors are driving these averages. To assess this, the following variables will be found regarding the sampled woodlot; area (ha), average slope (%), Length/Width ratio, boundary complexity, extraction method, and soil type. The data analysed and recorded will then be subject to statistical and multivariate regression analyses.

Through completing these objectives, the forestry industry could benefit from additional values for use in harvesting planning, costing and in turn, economic analysis. Information around small-scale woodlots is becoming increasingly important, as outlined in the review of literature.

4 Method:

Throughout the review of literature, a study using a combination of ArcGIS and Google Earth to assess infrastructure was not found. Therefore, this study used a combination of methods found throughout studies on unproductive area, landing size and forest roads, as well as knowledge on analyses using ArcGIS. For this study, a small-scale woodlot was defined as 0 – 500 hectares.

Step 1: Sample size and distribution determination.

This section of the method followed (Petherick, 2014) relatively closely as elements of the study were the same. The first necessary step was to determine the theoretical sample size necessary to achieve a certain level of confidence and precision. This was achieved using Cochran’s formula.

$$n_0 = \frac{Z^2 * p * q}{e^2} \tag{1}$$

Where; n_0 = sample size, Z^2 is the abscissa of the normal curve that cuts off area (α) at the tails, e is the desired level of precision, p is the estimated proportion of a factor that is expected to be in the population, and q is equal to 1 minus p.

A Z value was found for varying confidence intervals using Z-tables. The Z values for varying levels of confidence are displayed in table 2 below.

Table 2: Respective Z values for varying levels of confidence.

Z-Values for varying confidence levels	
Confidence Level (%)	**Z-Value**
80	1.282
85	1.44
90	1.645
95	1.96
98	2.33
99	2.576

The p-value relates to the degree of variability. Due to the degree of variability within the population not being known at the beginning of the study, a value of 0.5 was used. 0.5 is the maximum variability within a population; it is therefore conservative.

Cochran’s formula was then used to evaluate the sample sizes necessary to achieve the varying confidence intervals displayed in table 2, along with varying levels of precision (e). The results are displayed in table 3 below.

Table 3: Necessary sample sizes to achieve varying levels of confidence and precision.

Sample size (n) given confidence level and precision level						
	Confidence Level (%)					
Precision Level (%)	**80**	**85**	**90**	**95**	**98**	**99**
2.5	657	829	1082	1537	2172	2654
5	164	207	271	384	543	664
7.5	73	92	120	171	241	295
10	41	52	68	96	136	166

Based on the values displayed in table 3 and the time constraints to this study, an initial confidence level of 95 % was chosen with a precision level of 10%. Therefore, a sample size of 96 was chosen to be taken across New Zealand in order to be confident that 95% of the respective infrastructure values are within 10% of the average values gained. The confidence and precision levels were recalculated at the end of the study, but this provided a basis number of samples to aim for and a minimum level of accuracy that would be achieved.

The area of woodlots within New Zealand is not evenly split throughout the nine wood supply regions. Therefore, to ensure that the study was representative of the actual spread of woodlots throughout the country, a weighted average was used to determine the proportion of the 96 samples to be taken within each region. The results of this weighted average are shown below in table 4.

Table 4: Number of samples to be taken from each wood supply region.

Wood Supply Region:	% of total woodlot area (ha):	# of samples:
Northland	12.0%	11
CNI	18.0%	17
East Coast	6.2%	6
Hawkes Bay	7.2%	7
SNI	19.2%	18
Nelson/Marlborough	11.5%	11
West Coast	0.8%	1
Canterbury	10.8%	10
Otago/Southland	14.5%	14
Total	100%	96

Step 2: Preparation of ArcGIS files.

The shapefile containing the woodlots mapped by the BForSc students was provided by Vega Xu, Bruce Manley and the University of Canterbury. Figure 1 on the following page was created to display the spread of woodlots. As can be seen, it does not include Auckland, which is a part of the Northland wood supply region. Auckland was omitted from the study due to delays in the mapping of the Northland woodlots by the 2020 BForSc class; these delays occurred in response to the COVID-19 pandemic. Auckland is a very small area, and samples of woodlots were taken just outside of the Auckland boundaries; therefore, it was deemed that the exclusion of Auckland would not significantly impact the quality of the study. The shapefile needed to be refined in ArcGIS to contain the information wanted. The shapefile categorised woodlots into a status of "Forest", "Awaiting Restock" or "Windthrow". The category "Awaiting restock" represented harvested woodlots that were awaiting replanting at the time of analysis, therefore, were determined to be a good representation of woodlots that have been harvested and could be evaluated for their infrastructure. The woodlots were also categorised by region. The "Select by attributes" function in ArcGIS was used to select all woodlots that are "Awaiting restock", these were then exported into a new shapefile containing only forests that were "Awaiting restock". The function was then used again to select and export the awaiting restock woodlots into individual wood supply regions. The following table shows the number of forests that were present in each wood supply region that were awaiting restock.

Table 5: Number of woodlots awaiting restock that were in each region's woodlot shapefile.

Wood Supply Region:	**Number awaiting restock:**
East Coast	216
Hawkes Bay	180
SNI	333
Nelson/Marlborough	326
West Coast	74
Canterbury	535
Otago/Southland	914
Total	2578

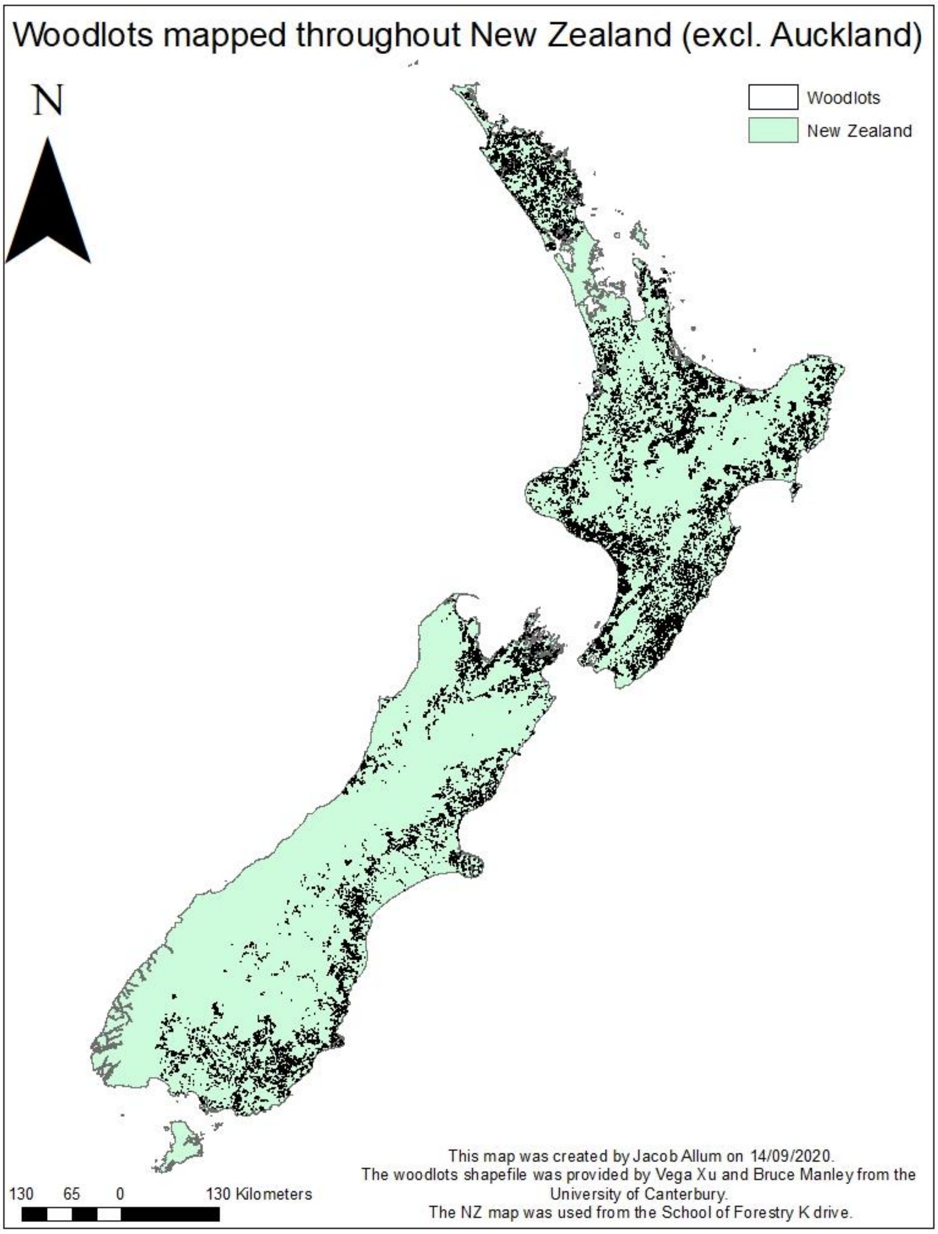

Figure 1: Spread of woodlots throughout NZ, excluding CNI and Northland.

Once the shapefiles for woodlots awaiting restock in each region were produced, a slope map of each region was required. To achieve this, an 8-metre digital elevation model (DEM) was downloaded from Land Information New Zealand's (LINZ) Data Service (https://data.linz.govt.nz/) for both the North and South Islands. The DEM's downloaded from LINZ Data Service in the form of 115 panels. The "Mosaic to New Raster" tool within ArcGIS was then used to merge the respective North Island panels into a North Island DEM and the respective South Island panels into a South Island DEM. The DEM for each island was then used as an input to the ArcGIS' "slope" function; the slope was defined in terms of percentage increase as opposed to degrees. Through this, a slope map for each island was produced.

The North Island slope map, which was generated following the above processes, is shown in figure 2.

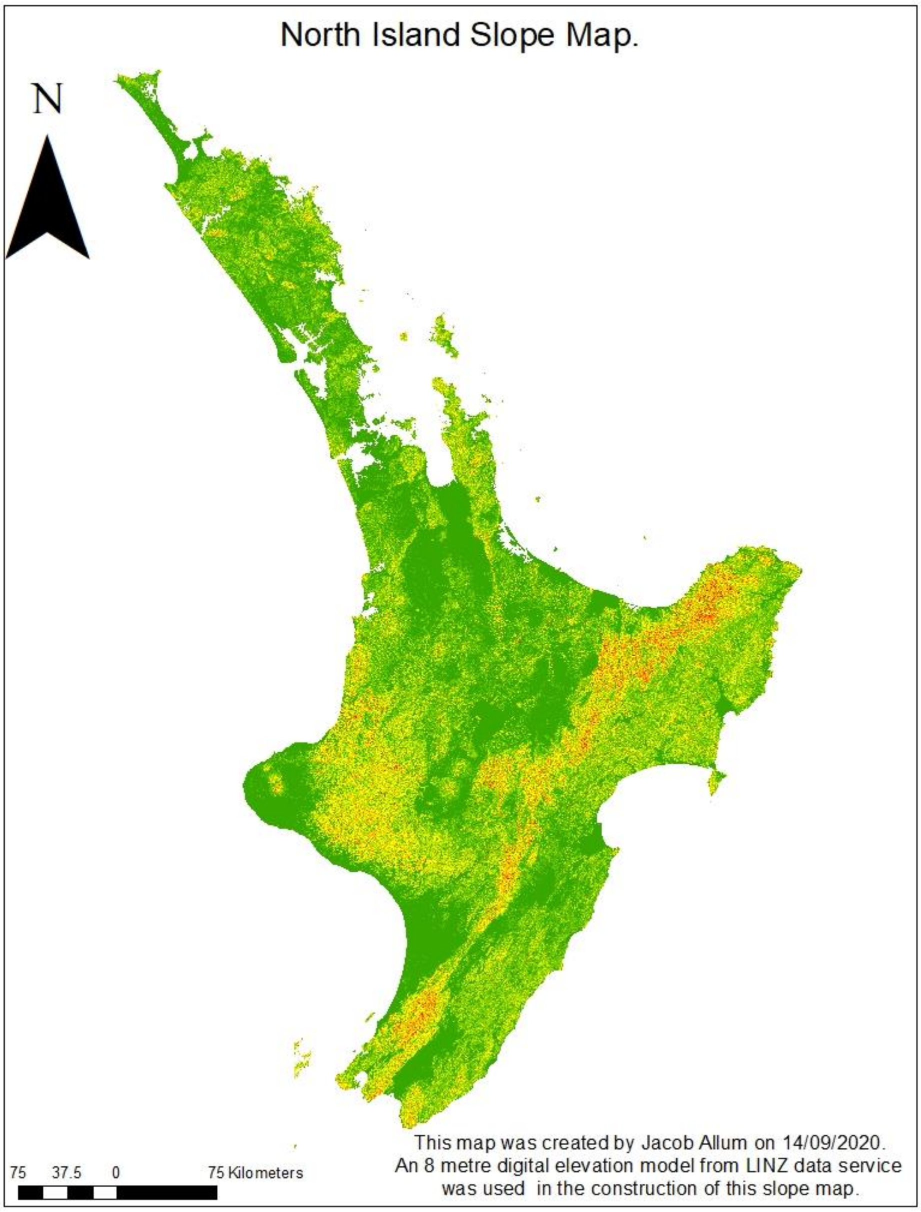

Figure 2: Slope map of the North Island.

Step 3: Random Sampling of Woodlots.

It would have been physically infeasible to evaluate all woodlots across New Zealand within the scope of this study. Therefore, as discussed in step 1, a sample size was chosen to achieve a theoretical confidence interval of 95 $\pm$ 10 %. The woodlots were then randomly sampled to acquire the number of samples for each region displayed in table 4. The number of samples taken within each region were then split between the size classes listed in the National Exotic Forest Description (NEFD) (Ministry for Primary Industries (MPI), 2019) up to 500 hectares. The classes were.

- Less than 40 hectares.
- 40 - 99 hectares.
- 100 – 499 hectares.

The proportion of each regional sample taken within each of the size classes was calculated using a weighted average based upon the proportion of total woodlot area within each size class. These are displayed in table 6 below.

Table 6: Number of samples taken within each wood supply region within each size class.

Wood Supply Region:	Total # of samples:	# of Samples 0 - 40 ha:	# of Samples 40-99 ha:	# of Samples 100 - 499 ha:
Northland	11	7	1	3
CNI	17	10	2	5
East Coast	7	4	1	2
Hawkes Bay	7	4	1	2
SNI	18	11	2	5
Nelson/Marlb	11	7	1	3
West Coast	1	1	0	0
Canterbury	10	6	1	3
Otago/South	14	8	2	4
Total	96	59	11	27

In the case where there were not enough woodlots to be sampled within a respective size class and region, the sample was taken from the size class below the one lacking enough samples. This only occurred on one occasion when only two samples of adequate visibility and clarity for assessment were found in the 100 – 499 ha size class for the Canterbury region. This meant that the 3rd sample that was meant to be taken from the 100 – 499 ha size class was taken from the higher end of the lower size class, 40 – 99 ha. This woodlot was equal to 91.28 hectares in size.

The randomly sampled woodlots were selected from within each regions shapefile of forests awaiting restock. Each regions sample was then exported to its own sample shapefile. At this point the sample shapefile was opened in Google Earth through selecting "open" within google earth, changing the file type from a Google Earth file (.kml) to an ESRI shapefile (.shp), and then selecting the saved sample shapefile for the region. Each woodlot mapped within the shapefiles included information in the attribute table regarding what year imagery was used when mapping was undertaken. This information was used to assist in determining what year of image to display in Google Earth using Google Earths built in historical imagery tool. Once opened in Google Earth each woodlot was checked to ensure a necessary level of clarity at the respective year that shows the woodlot in awaiting restock status. If visibility was deemed too low due to shadow or cloud cover or if the quality of image was not adequate to distinguish between relevant features, the woodlot was resampled for another woodlot within the same size class. The outline of the woodlot area was also assessed to check it follows the actual boundary of the forest. If for the given imagery year the outline of the woodlot seemed to not follow the woodlot boundary accurately, historical images were used to assess whether this was likely due to an error in mapping or whether satellite imagery angles were having an influence on where the outline is located. If it was determined that the area mapped was not similar to the true area, the woodlot was resampled. However, as outlined in the review of literature all woodlot mapping was peer reviewed and checked by experienced post graduate students so most gave an accurate representation of area.

Step 4: Data collection.

Once random sampling was completed, and the necessary image quality was confirmed the data collection step commenced. Firstly, each region was assigned a sheet in an excel spreadsheet. Each regional sheet contained two tables. One table included the infrastructure data for all samples within the region, the second included the variable data. The infrastructure data that was calculated was road density (m/ha) and landing size (m^2), along with the average area that each of these landings serviced (ha/landing). The variables that were recorded to calculate road density were road length (m), and woodlot area (ha). Due to most roading contractors being able to construct a forest road based upon a line plotted on a map (Brown & Visser, 2018) it was deemed unnecessary to record road width as a variable for necessary infrastructure, it would have also been very difficult to determine the edge of forest roads within Google Earth. The variables that were chosen to evaluate their impact on infrastructure were area (ha), average slope (%), length/width ratio, boundary complexity, soil type and extraction method. Throughout studying relevant literature these variables were noted to potentially impact infrastructure. The processes for calculating these variables is outlined later in the methodology. The additional variables that were required to be recorded in order to calculate these variables were number of landings, perimeter (m), longest continuous length (m), and longest continuous width perpendicular to the axis of length (m). Unfortunately, due to the size of this study and the nature of it being done remotely there were certain variables that were outlined in literature as potential effects that could not be recorded. This is a limitation of this study. Some of the variables that would have been beneficial to record but were not are number of log sorts, type of loader, and daily productivity.

Firstly, within ArcGIS the sample shapefile for a respective region was overlay on the slope map for its respective island. Then using the “zonal statistics as table” function in ArcGIS, with the slope map as the input raster and a region’s sample shapefile as the input polygons, the average slope (%) was calculated over the respective polygon areas and recorded in the excel spreadsheet. An example of a zonal statistics as table output for the Nelson/Marlborough sample is shown below in table 7. The focus of the overall output is on the mean output, however, the other results from the output have still been included.

Table 7: Zonal Statistics as Table output for the Nelson and Marlborough sample.

FID	COUNT	AREA	MIN	MAX	RANGE	MEAN	STD	SUM
0	532	34048	8.1	47.8	39.7	29.1	8.0	15496.2
1	3729	238656	2.8	132.2	129.4	69.2	23.4	258108.8
2	1489	95296	37.1	126.7	89.6	75.8	17.0	112850.5
3	3297	211008	0.4	77.0	76.6	40.8	16.4	134413.5
4	904	57856	2.6	112.2	109.6	57.4	23.0	51894.9
5	2543	162752	17.1	165.2	148.1	77.4	24.5	196917.7
6	6092	389888	1.3	102.8	101.5	50.7	17.6	309089.0
7	15192	972288	0.1	207.2	207.0	68.7	30.0	1043886.0
8	16739	1071296	3.8	112.4	108.5	66.3	13.1	1109565.4
9	27518	1761152	21.9	182.9	161.0	86.3	21.6	2374013.8
10	23928	1531392	0.0	22.7	22.7	1.9	1.5	44275.1

The Nelson/Marlborough sample overlaid on the South Island slope map to achieve the above results, is shown below in figure 3. The black outlines represent the woodlots in the sample.

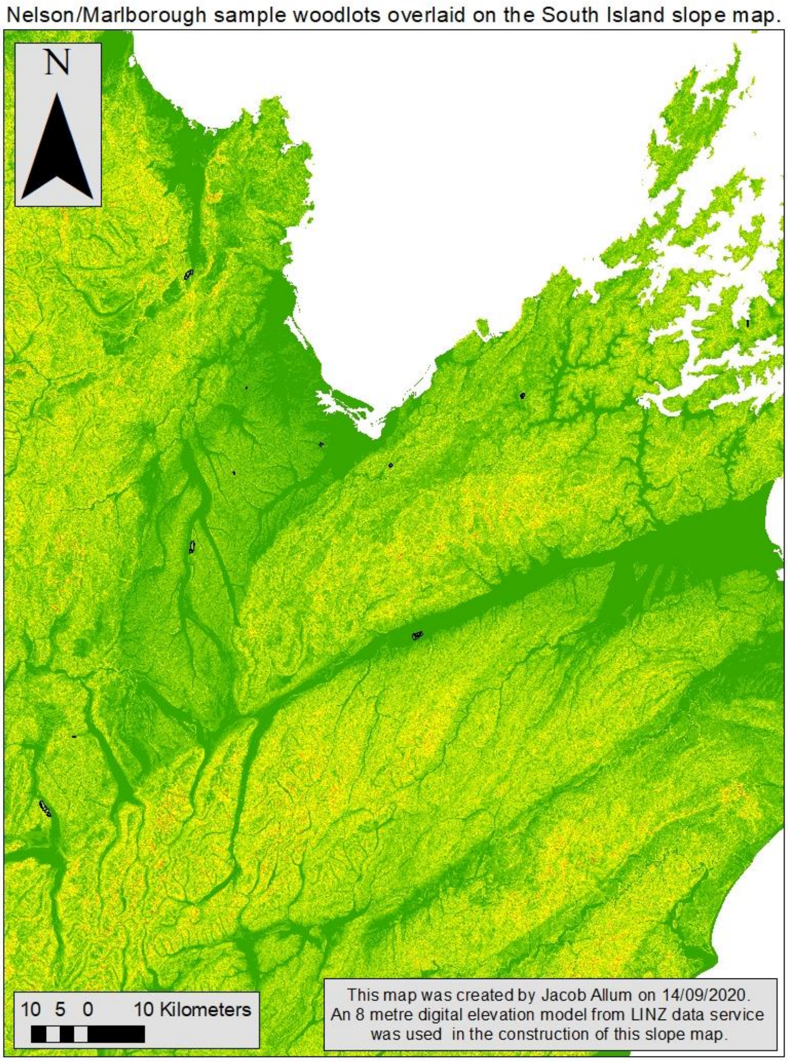

Figure 3: Example woodlot outline overlaid on the South Island slope map.

Within ArcGIS, prior to transferring the analysis to Google Earth, the areas (ha) of the respective sampled woodlots were found in the attribute table of the sample shapefile. These were then recorded in the excel spreadsheet.

The perimeters were also calculated in ArcGIS. The attribute table of the respective shapefile was opened and the option to add a field was selected, this was then named perimeter (m). The field output was selected as type ‘float’. Then, after opening an edit session (to allow for changes to be made if a mistake is made), the ‘calculate geometry’ feature was used. Within the ‘calculate geometry’ feature the output was selected as perimeter, based upon the NZ transverse Mercator coordinate system, with the units as metres. Once this was completed the respective perimeters for each of the sampled woodlots were recorded in the excel spreadsheet.

Next the analysis was transferred to Google Earth. As outlined in the review of literature, studies have quantified the error in measurements obtained from Google Earth as relatively low. Therefore, it was deemed as an easy solution to calculating the remaining variables needed. For all measurements and plotting undertaken in Google Earth, professional judgement was used to determine the exact location measurements and plots should end. To assist in the professional judgement, historical imagery was assessed to see differences between years, also once a measurement was taken or road/landing plotted, a full 360° view of the respective measurement/plot was evaluated to ensure no errors were made based upon viewing angle. Measurements and plotting were completed at the closest possible zoom that still provides clear imagery.

Firstly, using the line measuring function within Google Earth, the longest, continuous length of the woodlot was measured and recorded as the length (m) in the excel spreadsheet. The angle that the length was plotted at was also recorded. The largest width was then found at a perpendicular angle to the longest length, this was recorded as width (m) in the data spreadsheet. An example of this process for an example woodlot in the Nelson/Marlborough region is shown below in figures 4 and 5. The yellow lines represent the measurement being taken. As can be seen the length measurement was taken at an angle of 352.22°, therefore, the width measurement was taken at a bearing of 82.22°, a 270° difference.

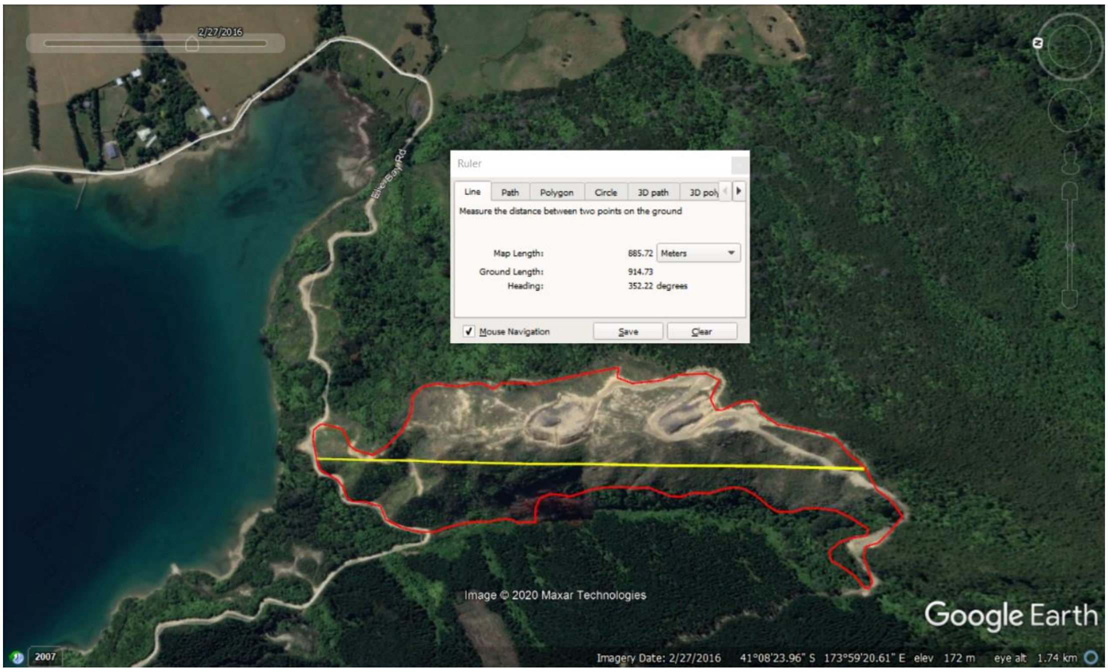

Figure 4: Measuring length of the example woodlot.

Figure 5: Measuring width of the example woodlot.

The number of landings within each woodlot were then counted and recorded in the excel spreadsheet, being assigned a number at the same time. A landing was defined as any compacted, flat, contiguous area that has been clearly used as a landing.

The area of each numbered landing was then found by using the measurement tools within Google Earth to plot a polygon shape around the landing area, this plot was then labelled in relation to the numbered landing and saved. The area of this polygon was then recorded in the data collection spreadsheet. If a road ran through a landing it was included in the area, if a road ran alongside a landing it was excluded. In the case that the boundary between the landing edge was blurred due to erosion or slash, historical imagery was used to attempt to determine the landing edge with more clarity, as well as a 360° assessment. In a couple of cases the edge of the landing was still not 100% clear but historical imagery and a 360° assessment gave a relatively accurate idea of where the landing ended. In these cases, a conservative plot was completed based upon professional judgement.

Using the ‘path’ function within the measuring tools in Google Earth the roading within the woodlots was then plotted. When plotting the path, the plots were made as close to the centre point of the road as possible, this was meant to represent the line plot of a road that contractors would be able to compete a design from. Any roading that runs into subsequent lots or outside of the forest was not included. Two exceptions to this rule were; if the road briefly ran outside and back inside of the plotted boundary for a switchback, or if a road ran along a boundary and was deemed to be a necessary construction to harvest the woodlot. Roads ran from the forest boundary and stopped at the landing edge. Any roading within the boundaries of a landing was considered as landing area. The path was then saved as the road for the woodlot. The length of the path was then recorded in the data collection spreadsheet as the road length (m) for the respective woodlot.

An example of landing and road plots for the same example woodlot in the Nelson/Marlborough region are shown below in figure 6. The pop-up box in the figure shows the landing area and perimeter for the first mapped landing. The outer ring of the selected landing was determined as a slash bench.

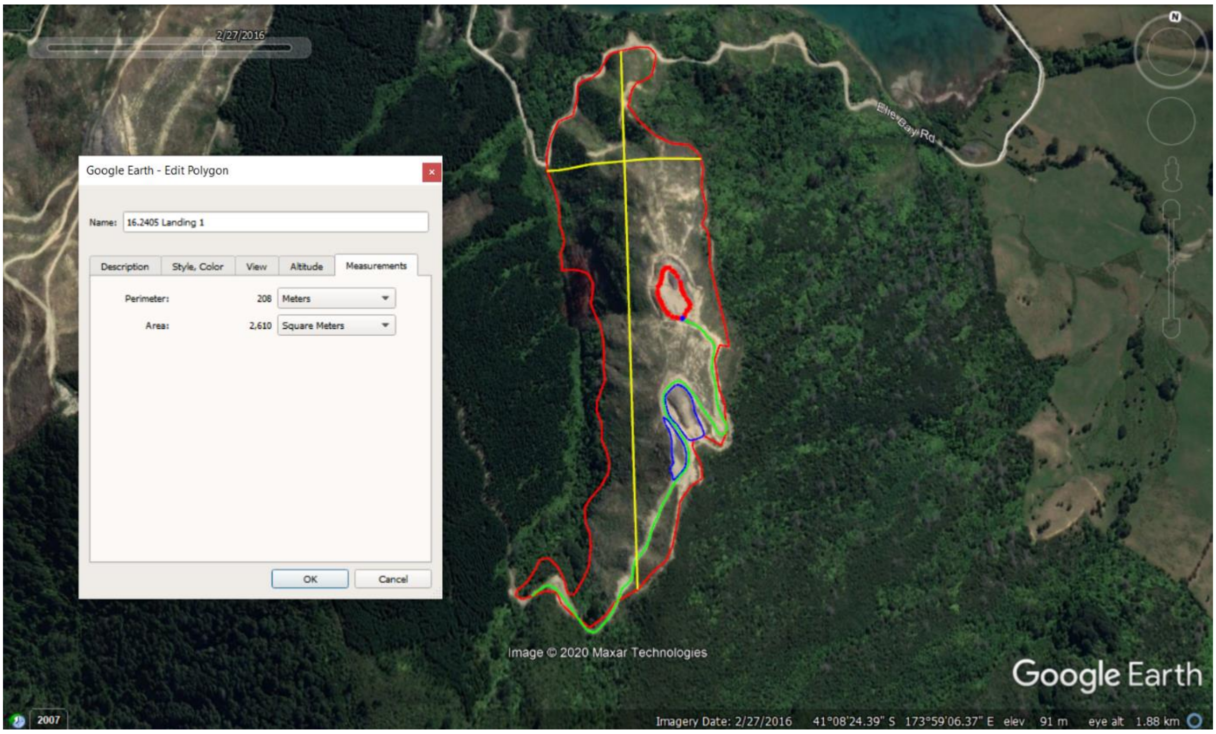

Figure 6: Plots of landings and roading within an example woodlot.

The primary extraction method used for harvesting the woodlot was defined as the extraction method used for greater than 50% of the woodlot area. The primary extraction method was determined based on professional judgement and taking into account several variables. The primary variable used was the slope map. Through using 40% as an upper limit for ground-based harvesting it was assumed that a hauler extraction method was used for slopes greater than 40%. However, another variable taken into consideration was the layout of skid trails and roading. Also, the area surrounding landings was assessed for scarring to determine a pull direction, and hence a hauler system. The determined extraction method was then recorded in the data spreadsheet. In the case where the extraction methods were deemed to be used around 50% each the extraction method was recorded as “both”.

To evaluate the soil type within the woodlot the location was cross-referenced with the online New Zealand soil classification map provided by Manaaki Whenua - Landcare Research (Manaaki Whenua - Landcare Research, 2020). An example of this process is shown in the following figures. The location was assessed using Google Earth and then the corresponding location was found on the New Zealand Soil Classification Map. The location of the woodlot was then clicked, and the soil classification was given, this is shown in figures 7 - 9. The soil type was then recorded for that woodlot.

Figure 7: Location of the woodlot on Google Earth.

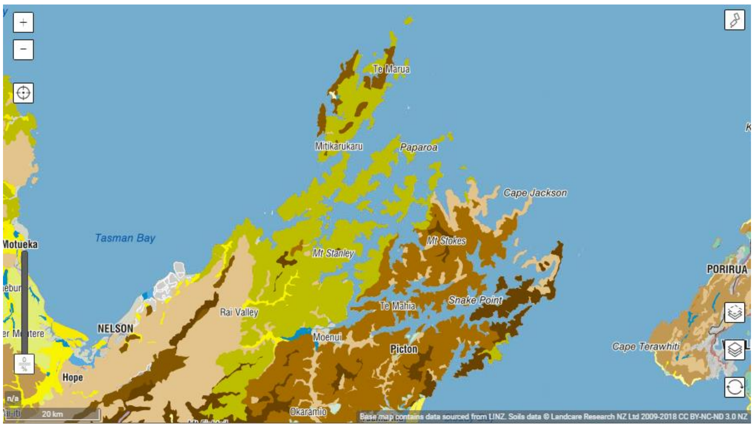

Figure 8: Soil classifications over a wider region.

Figure 9: Same view as the Google Earth imagery in figure 6, allowing soil classification.

The steps outlined throughout step 4 show how the relevant data was found for one woodlot. This process was then repeated for the entire sample within a region. Once a region was complete, the analysis was moved to another region.

Step 5: Data processing and analysis.

Once the samples for each region were completed, data manipulation was required to calculate more influencing variables, road density and landing service area. Firstly, the road density of each woodlot was calculated in excel. This was done using the following equation.

$$Road\ Density = \frac{Road\ Length\ (m)}{Woodlot\ Area\ (ha)} \tag{2}$$

The average landing service area for a woodlot was then calculated using the following equation.

$$Avg\ Landing\ Service\ Area = \frac{Woodlot\ Area\ (ha)}{\#\ of\ Landings} \tag{3}$$

The variables length/width ratio and boundary complexity, that were assessed for their influence on the infrastructure, were calculated using equations 4 and 5.

$$\frac{Length}{Width} Ratio = \frac{Length\ of\ the\ woodlot\ (m)}{Width\ of\ the\ woodlot\ (m)} \tag{4}$$

$$Boundary\ Complexity = \frac{Area\ of\ woodlot\ (m^2)}{Perimeter\ of\ woodlot\ (m)} \tag{5}$$

The average values for infrastructure were first computed by region, and then nationally. To find the average values the following equations were used.

$$Avg. landing\ Size = \frac{\sum Landing\ areas\ (m^2)}{Number\ of\ landings} \tag{6}$$

$$Avg. road\ density = \frac{\sum Road\ densities\ (\frac{m}{ha})}{Number\ of\ samples} \tag{7}$$

$$Avg.\ landing\ service\ area = \frac{\sum Avg\ Landing\ Service\ Area}{Number\ of\ landings} \tag{8}$$

Average values for each of the infrastructure variables were also calculated for each extraction method (hauler or ground-based) and differing soil classifications.

Once average values were computed a statistical analysis was conducted. It compared the mean and percentiles between soil type and extraction methods in order to determine any significant difference between these variables.

A multivariate regression analysis was then completed in excel on each of the infrastructure variables. This enabled an assessment of how the dependent variables (average infrastructure values) changed when one of the recorded independent variables (influencing data) changed. This allowed the determination of which variables influence the infrastructure averages gained throughout the study. Through this, an assessment of how well the regression equations predicted the dependent variables and the level of influence each variable had was evaluated.

5 Results and Analyses:

5.1 Average infrastructure values:

Throughout this project 96 woodlots across New Zealand were evaluated to find three variables relating to infrastructure: the average road density, the average landing size, and the average service area of each landing.

After recording and calculating the necessary values to gain all infrastructure averages the results were displayed in table 9 below.

There were a total of 15 woodlots that had no roading within the boundaries of the woodlot. Whether these 15 woodlots are included in the national average road density calculation has a significant impact on the value gained. When these woodlots are included the average road density is 25.2 m/ha, when they are not included in the calculation the average road density is 29.9 m/ha, a 4.7 m/ha difference. Therefore, both values have been included in the results summary table below. In the table, road density is recorded as R.D, and landing service area is recorded as L.S.A.

Table 8: Average infrastructure values calculated by region and nationally.

Region	# of Samples Taken	R.D (m/ha) (All)	R.D (m/ha) (No 0m/ha)	Landing Size (m2)	L.S.A (ha/landing)
Northland	11	24.0	24.0	3359.0	14.5
CNI	17	25.2	26.7	3283.2	11.7
East Coast	7	19.1	26.8	2746.6	10.7
H. Bay	7	25.6	35.8	2785.1	7.9
SNI	18	27.5	33.0	2661.2	8.2
Nels/Marl	11	35.0	42.8	3120.6	15.5
West Coast	1	33.3	33.3	2134.3	5.3
Canterbury	10	18.0	25.8	2748.7	18.1
Otag/South	14	22.8	26.6	3191.3	16.9
NZ	**96**	**25.2**	**29.9**	**3000.1**	**12.8**

The sample size of 96 was chosen to achieve a minimum level of confidence and precision of 95 +/- 10% nationally, based upon Cochran's formula. This same level of confidence and precision only applies nationally. Therefore, theoretically, based upon Cochran's formula, we can say with 95% confidence that the average values of road density, landing size, and average landing service area are within +/- 10% of the following values for woodlots up to 500 hectares in size within New Zealand;

- Road Density (All Lots Included) = 25.2 m/ha (1dp)
- Road Density (No 0m/ha lots included) = 29.9 m/ha (1dp)
- Landing size = 3000.1 m^2 (1dp)
- Landing Service Area = 12.8 ha/landing (1dp)

However, Cochran's formula is theoretical, and the true accuracy of the infrastructure values varies slightly from this confidence and precision level. Using excels descriptive statistics function the following ranges for a 95% confidence interval were calculated for the respective infrastructure values. These are displayed below in table 9.

Table 9: Experimental confidence intervals for the calculated infrastructure values.

Infrastructure Variable	95% Confidence Interval
Road Density (All Lots Included)	25.2 ± 3.6 m/ha
Road Density (No 0m/ha lots included)	29.9 ± 3.4 m/ha
Landing Size	3000.1 ± 230.0 m^2
Landing Service Area	12.8 ± 1.8 ha/landing

For reference, table 10 below shows the average values calculated for all the numerical independent variables (no extraction method or soil type included) that will be used to assess their influence on the above averages. Boundary complexity is shortened to B.C in table 10.

Table 10: Average values for the recorded independent variables.

Region	# of Samples Taken	Area (ha)	Slope (%)	L/W Ratio	B.C
Northland	11	56.2	37.9	1.8	73.3
CNI	17	68.9	40.7	2.7	103.6
East Coast	7	87.0	50.4	2.2	85.3
H. Bay	7	49.9	44.8	4.6	62.5
SNI	18	56.2	47.6	1.8	92.2
Nels/Marl	11	59.3	56.7	2.4	97.5
West Coast	1	16.0	4.1	2.7	72.0
Canterbury	10	76.5	32.9	1.9	113.5
Otag/South	14	90.0	25.2	2.5	100.8
NZ	**96**	**67.2**	**41.1**	**2.4**	**93.3**

5.2 Distributions:

The distributions of all the infrastructure data from each woodlot were plotted in histograms in order to see the spread of results and whether the data conforms to a normal distribution. Figure 10 and 11 below show the distribution of road densities calculated throughout the 96 woodlots across New Zealand. Figure 10 shows the distribution with the 0 metre/hectare values included, figure 11 shows the distribution with these 0 metre/hectare results removed.

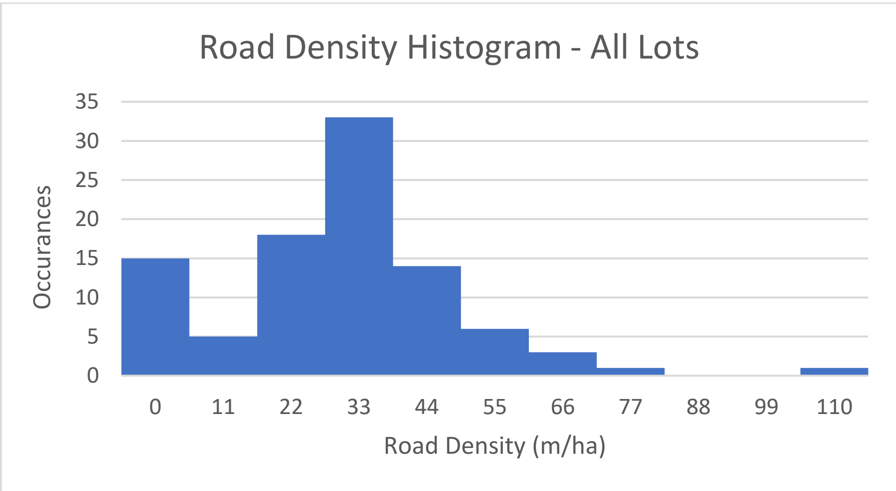

Figure 10: Histogram showing the distribution of Road Density results (incl. 0 m/ha values).

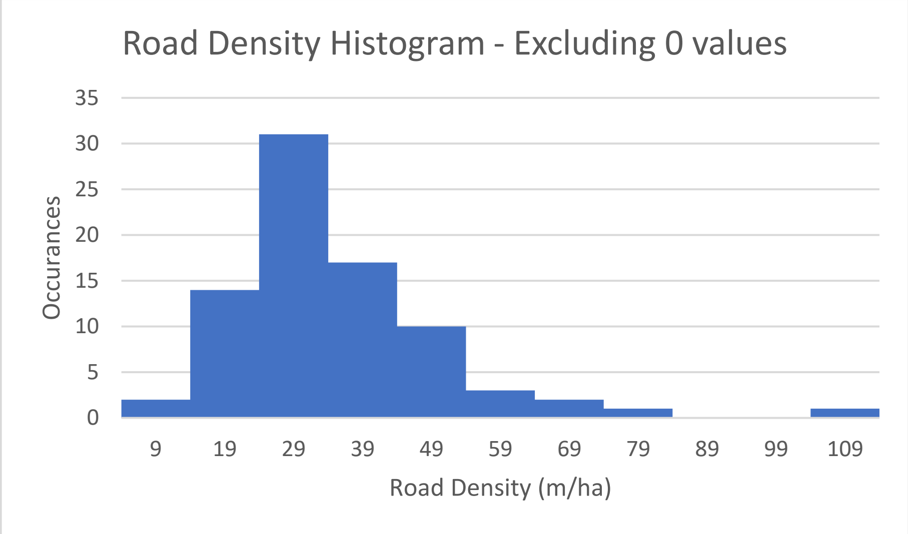

Figure 11: Histogram showing the distribution of Road Density results (excl. 0 m/ha values)

As can be seen from figures 10 and 11 the distribution of road density results shows a relatively normal distribution which is skewed to the right. If the outlier of 105.4 m/ha was removed from the results, then the distribution would still be slightly skewed to the right but to a lesser extent.

A histogram showing the distribution of landing size results from the 96 woodlots is shown below in figure 12.

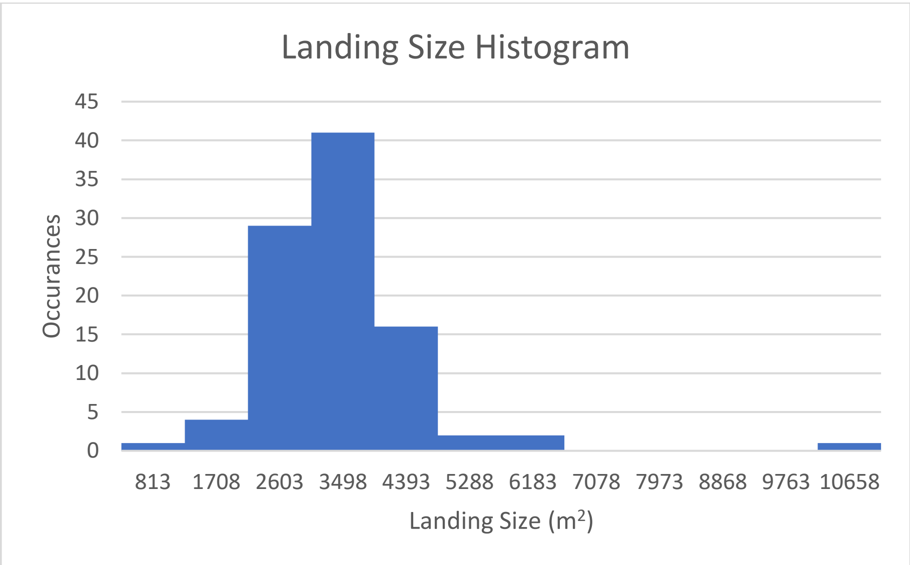

Figure 12: Histogram showing the distribution of landing size results.

Figure 12 also shows a normal distribution that is skewed to the right. However, if the outlier of 9763.67 m^2 was removed from the histogram, the rightwards skew would be removed and we would see a normal distribution with no skew. It was expected that landing size produced the most normal distribution as a much higher number of samples were taken for landing size overall, as most of the time multiple landings were recorded within each sampled woodlot.

The final histogram showing the distribution of average landing service area results is displayed in figure 13 below.

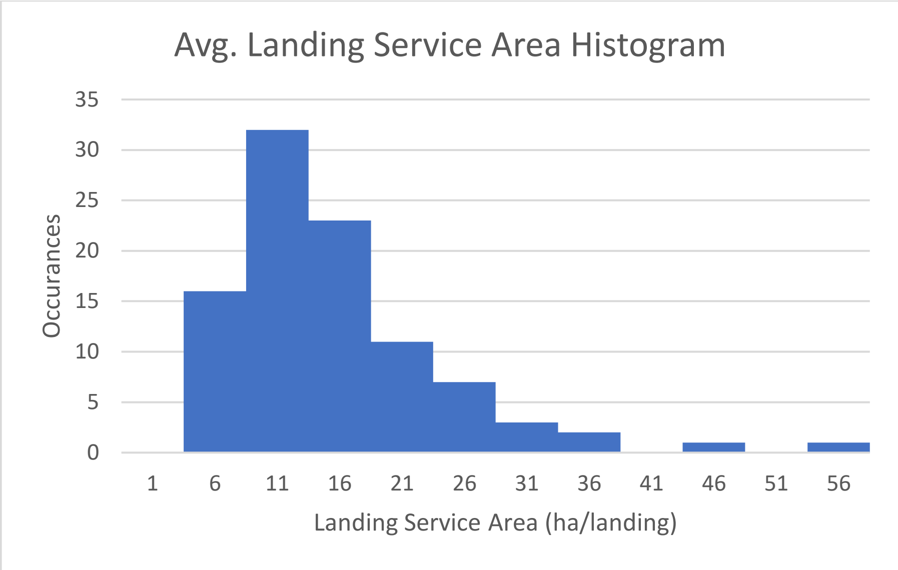

Figure 13: Histogram showing the distribution of average landing service area results.

The average landing service area distribution shows a relatively normal distribution that is heavily skewed to the right. The two high values in the 41 – 46 and 51 – 56 ha/landing bins exaggerate the righthand skew, however, even if these were ruled as outliers there would still be a heavy right-hand skew of these results.

5.3 Regression Analyses and Line Fit Plots:

To assess which of the independent variables have an influence on the dependent variables multivariate regression analyses were run against the four infrastructure values. Along with this, individual line fit plots were produced for the variables that were deemed to have a significant influence. All regression analyses outputs are displayed in Appendix I at the back of this report.

Firstly, a multivariate regression analysis was run for the four numerical independent variables (area, average slope, length/width Ratio, and boundary complexity) against the road density output including all lots. The table labelled multivariate regression 1, shown in Appendix I, shows the output from this multivariate regression analysis. The analysis showed that including the 0 m/ha road density values in the analysis had a detrimental impact on the viability of the analysis. The first regression run produced an R-squared value of 0.08301, meaning that the model accounted for 8.301 % of the variability of the road density data, which is very low. The null and alternate hypotheses for the overall model were;

Null Hypothesis: H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (No useful relationships between the independent variables and the dependent variable).

Alternate Hypothesis: H_1: $\beta_1 / \beta_2 / \beta_3 / \beta_4 \neq 0$ (At least one useful relationship between an independent

variable and the dependent variable).
The significance F value was 0.09262, this is greater than the significance level of 0.05. This means that we cannot reject the null hypothesis. Therefore, the overall model test failed meaning the model needed to be reconsidered. The regression was re-run excluding the variables with the lowest significance but consistently performed poorly. Therefore, focus and attention in regard to regression analyses was given to the road density values with the 0 m/ha values excluded.

When the multivariate regression was run for the four independent variables against the road density values excluding the 0 m/ha samples, a much higher quality model was produced. The output from this regression analysis is labelled "multivariate regression 2" in Appendix I. The initial regression achieved an R-squared value of 0.2962 showing that this model accounts for 29.62 % of the variability of road density. The null hypothesis and alternate hypothesis were considered the same as that from multivariate regression 1 above. The significance F value was found to be 1.958×10^{-5}, this means that we accept the alternate hypothesis, meaning there is at least 1 relationship between the independent variables and dependent variable. The individual t-tests of the respective independent variables were then assessed. The null hypothesis and alternate hypothesis for the individual t-tests were as follows;
Null Hypothesis: H_0: $\beta_1 = 0$ (the independent variable has no relationship with the dependent variable).
Alternate Hypothesis: H_1: $\beta_1 \neq 0$ (the independent variable has a relationship with the dependent variable).
The p-values for each independent variable and whether the null hypotheses are rejected are shown in table 11 below.

Table 11: P-values for the four independent variables.

Independent Variable	P-value	Reject Null Hypothesis
Area (ha)	0.92	No
Average Slope (%)	< 0.001	Yes
Length/Width Ratio	< 0.001	Yes
Boundary Complexity	0.094	No

These p-values show that the length/width ratio is the most significant variable influencing road density, followed by average slope. The p-values for boundary complexity and area meant that the null hypothesis could not be rejected and a relationship between them and road density cannot be confirmed.

Because area was deemed by these values to have the lowest influence on road density, it was removed from the regression and the regression was re-run. The results from the re-run regression are labelled as "Multivariate regression 2 repeat excluding area". The R-squared value for this model was found to be 0.2961, very similar to that of the previous model. It also passed the overall test, accepting the alternate hypothesis that at least one of the independent variables has a relationship with road density. The p-values for the independent variables from the individual t-tests are shown in table 12 below.

Table 12: P-values for the four independent variables after excluding area from the regression.

Independent Variable	P-value	Reject Null Hypothesis
Average Slope (%)	< 0.001	Yes
Length/Width Ratio	< 0.001	Yes
Boundary Complexity	0.012	Yes

Once the regression was re-run with the exclusion of area from the model the influence of boundary complexity became significant as the p-value was less than 0.05, meaning we could reject the null hypothesis and accept the alternate hypothesis that boundary complexity does have a relationship with road density. Overall, based upon this multivariate regression analysis we can say that average slope, length/width ratio and boundary complexity all have a relationship with road density. Length/width ratio has the strongest relationship, followed by average slope, followed by boundary complexity.

Based upon these three independent variables having a significant influence on road density, line-fit plots were created to further explore the relationships with road density. The line-fit plots show the actual results compared to the predicted results based upon a linear regression. The line-fit plots for road density against, average slope, length/width ratio and boundary complexity are shown in figures 14, 15, and 16 below.

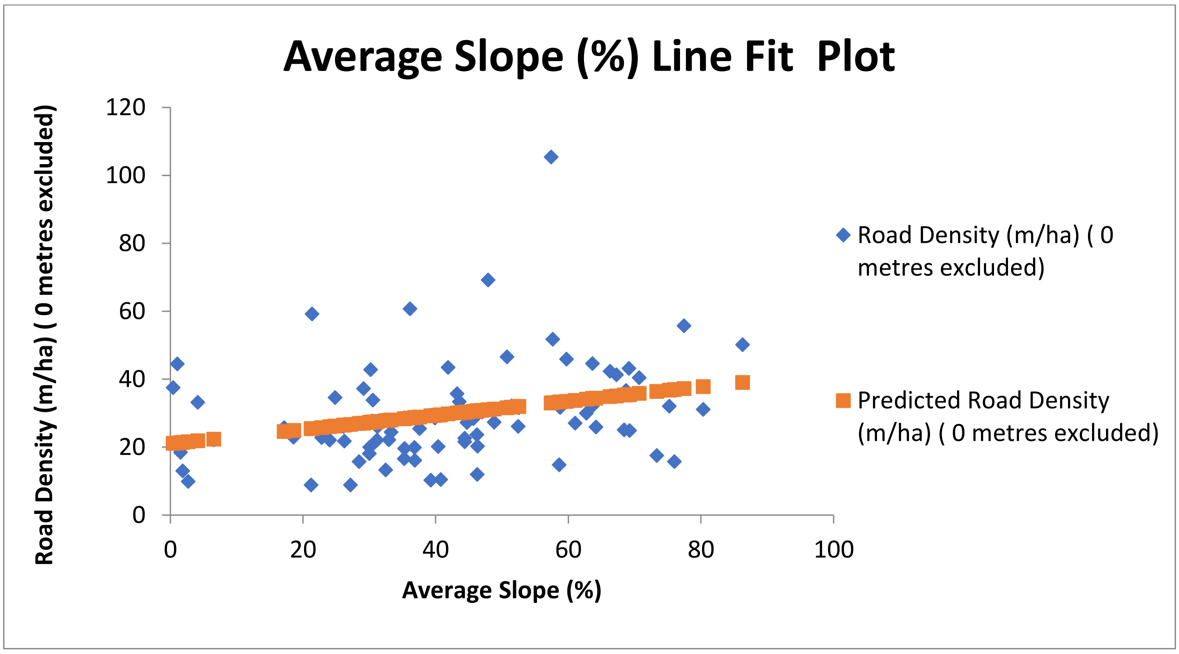

Figure 14: Line fit plot for road density against average slope.

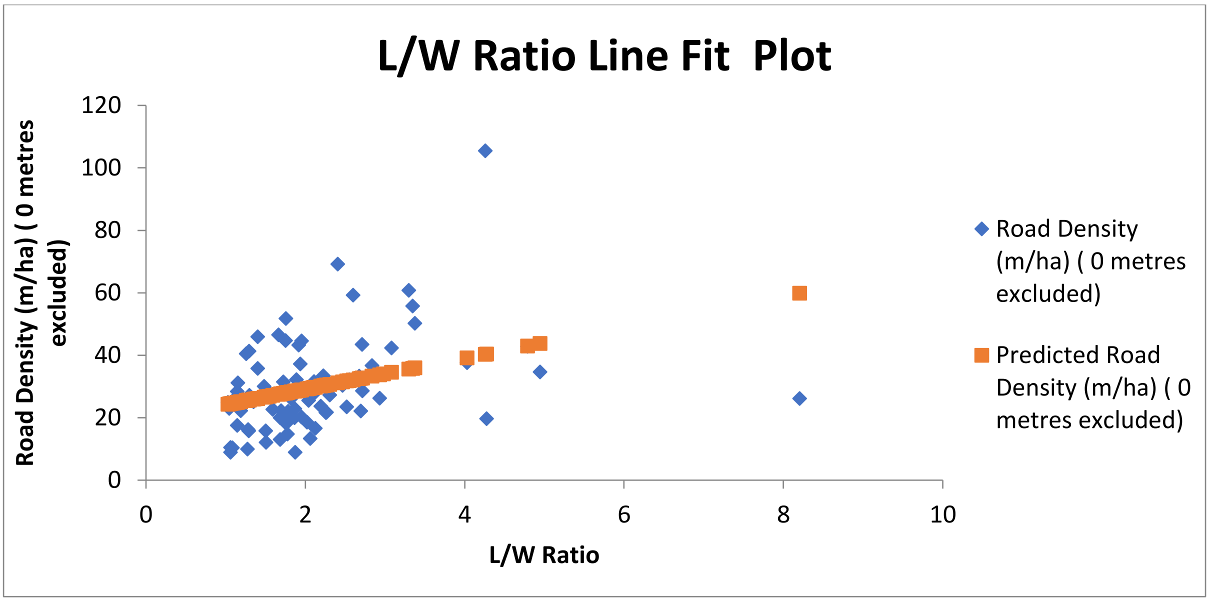

Figure 15: Line fit plot for road density against length/width ratio.

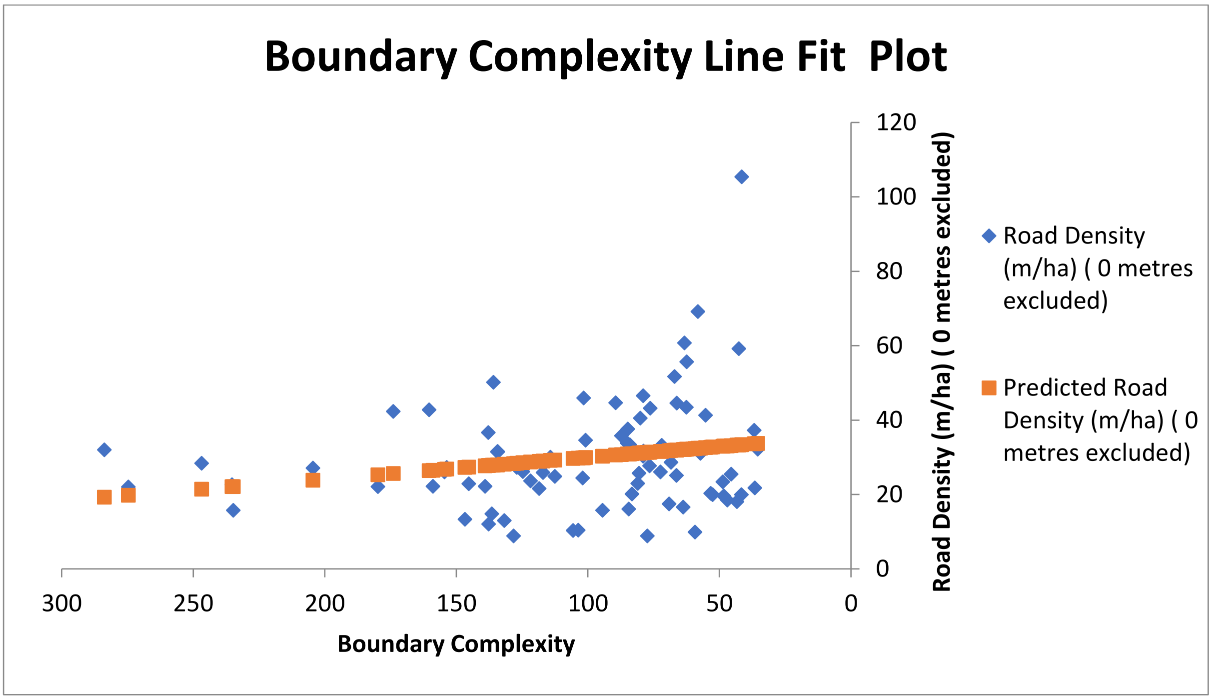

Figure 16: Line fit plot for road density against boundary complexity.

Next a multivariate regression analysis was run using the same independent variables, however, this time assessing their influence on landing size. The section labelled "multivariate regression 3" in Appendix I shows the output from this analysis. The R-squared value for the model was relatively low at 0.1936, meaning the model only accounts for 19.36% of the variability of landing size. Once again, the null and alternate hypotheses for the overall model were as follows;

Null Hypothesis: H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (No useful relationships between the independent variables and the dependent variable).

Alternate Hypothesis: H_1: $\beta_1 / \beta_2 / \beta_3 / \beta_4 \neq 0$ (At least one useful relationship between an independent variable and the dependent variable).
The significance F value gained was 0.0005498 and therefore the null hypothesis was rejected, and alternate hypothesis accepted. The null hypothesis and alternate hypothesis for the individual t-tests were again;
Null Hypothesis: H_0: $\beta_1 = 0$ (the independent variable has no relationship with the dependent variable).
Alternate Hypothesis: H_1: $\beta_1 \neq 0$ (the independent variable has a relationship with the dependent variable).
The individual p-values and whether the null hypothesis is rejected are shown in table 13 below.

Table 13: P-values for the four independent variables.

Independent Variable	P-value	Reject Null Hypothesis
Area (ha)	0.35	No
Average Slope (%)	< 0.001	Yes
Length/Width Ratio	0.028	Yes
Boundary Complexity	0.42	No

These values show that average slope has the strongest relationship with landing size, followed by length/width ratio. Due to the p-values being greater than 0.05 for area and boundary complexity the null hypotheses could not be rejected. Boundary complexity was deemed to have the weakest relationship with landing size and therefore was removed from the model before it was run again.

The re-run model output can be found in Appendix I under the label "Multivariate regression 3 repeat excluding boundary complexity". The R-squared value for this model was 0.1876, again a low R-squared value. It passed the overall model test with a significance F value of 0.0002461, meaning the null hypothesis was rejected and alternate hypothesis accepted. The individual p-values relating to the individual t-tests are shown below in table 14.

Table 14: P-values for the four independent variables after excluding boundary complexity from the regression.

Independent Variable	P-value	Reject Null Hypothesis
Area (ha)	0.022	Yes
Average Slope (%)	< 0.001	Yes
Length/Width Ratio	0.017	Yes

After re-running the model with the exclusion of boundary complexity the p-value for area was also less than the significance value of 0.05. That meant that now the null hypothesis could be rejected, and alternate hypothesis accepted, for all three independent variables. This means that area, average slope and length/width ratio all can be deemed to influence landing size. Average slope has by far the biggest influence on landing size, followed by length/width ratio, which is closely followed by area.

Line-fit plots were therefore developed for landing size against average slope, length/width ratio and area, in order to further explore the relationships. The line fit plots are displayed in figures 17, 18 and 19 below.

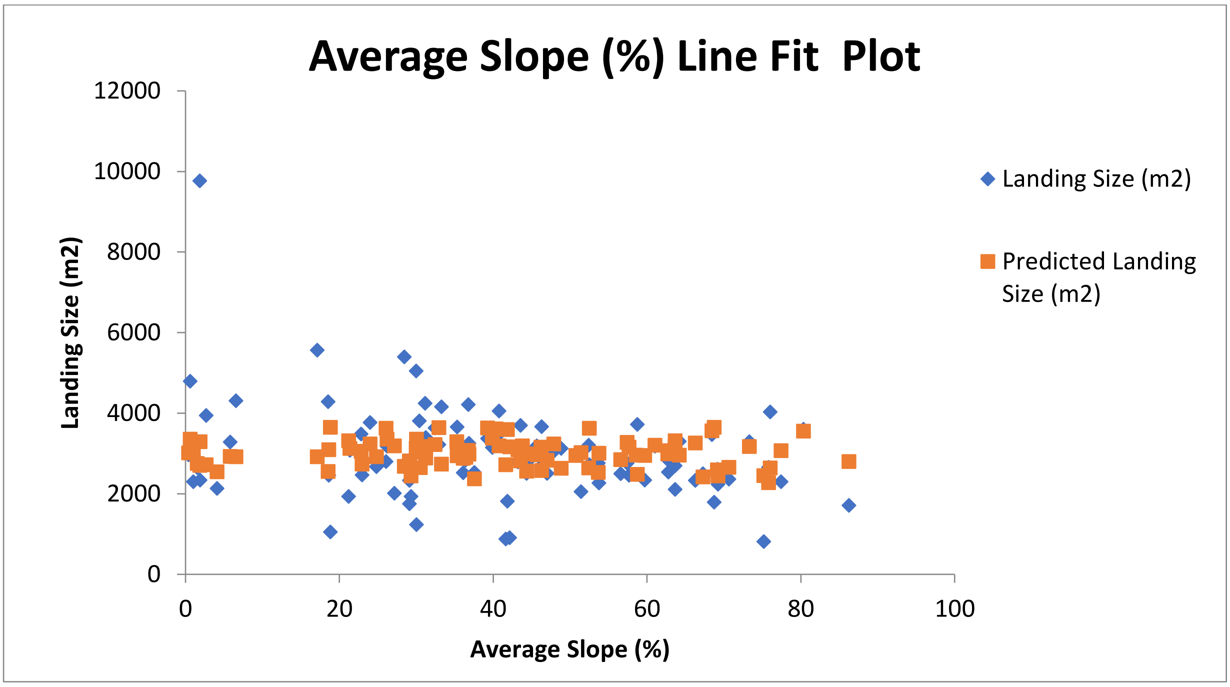

Figure 17: Line fit plot for landing size against average slope.

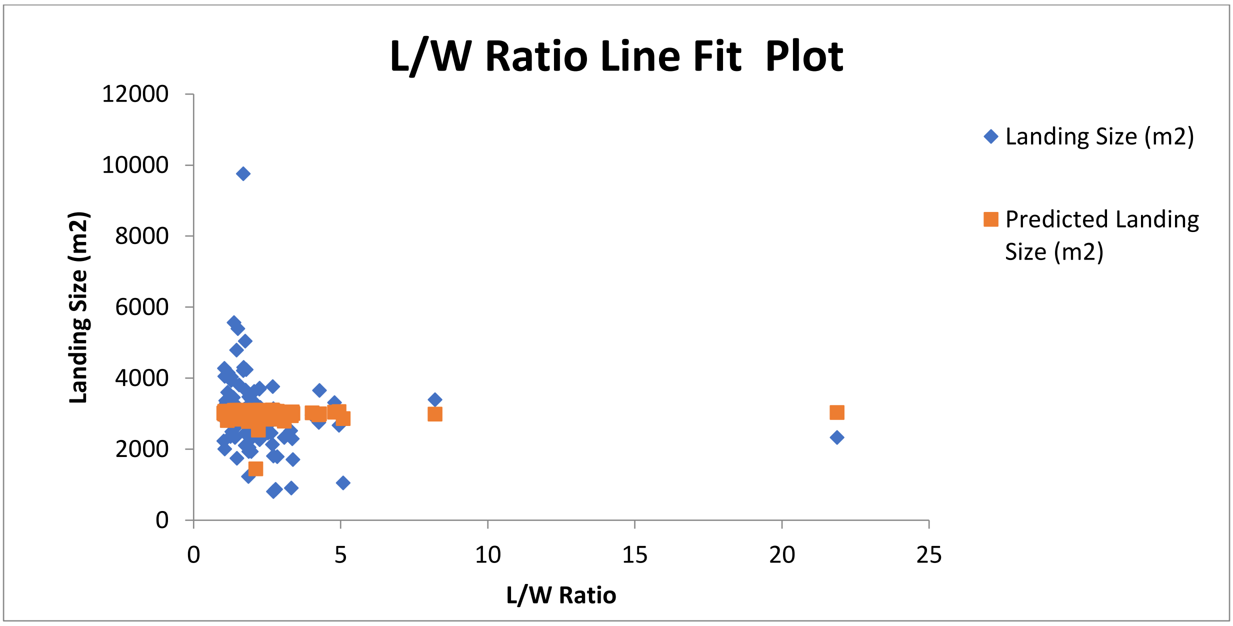

Figure 18: Line fit plot for landing size against length/width ratio.

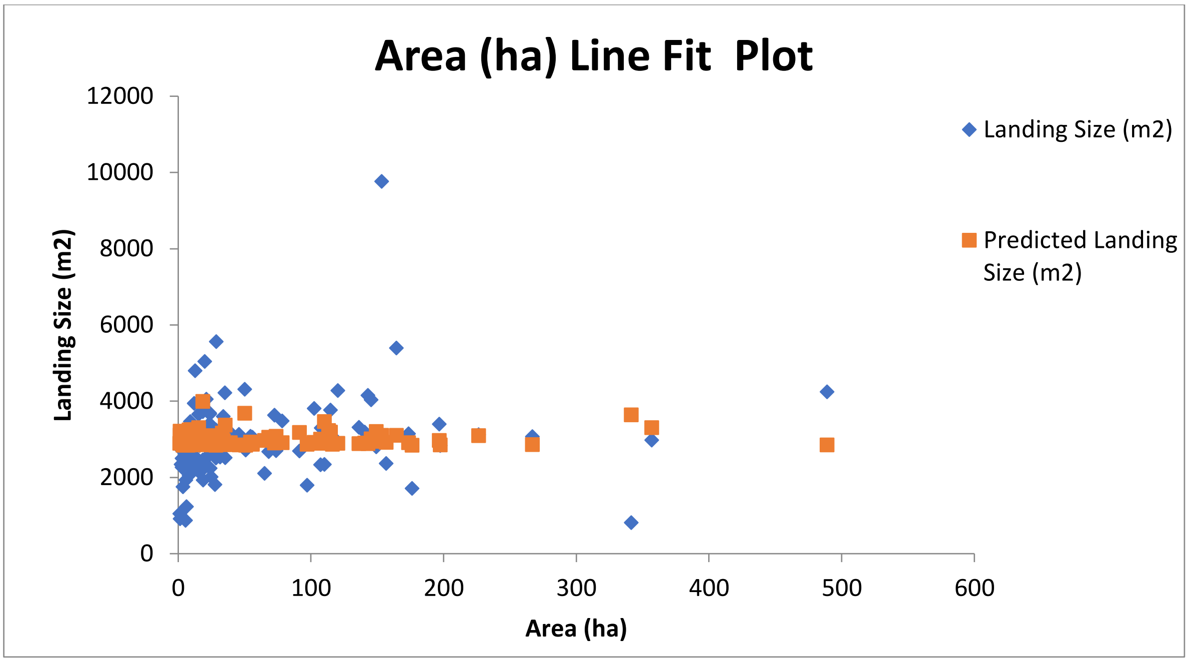

Figure 19: Line fit plot for landing size against area.

The final multivariate regression analysis run used the same independent variables and aimed to assess if they have a significant impact on the average area that landings within woodlots service. The outputs from this regression analysis can be found in Appendix I under the heading "Multivariate regression 4". This model was the strongest of all models run through a regression analysis, recording a R squared value of 0.3721, in other words accounting for 37.21% of the landing service areas variability. The null and alternate hypotheses for the model were kept the same as the previous multivariate regressions. A significance F value of 1.138×10^{-8} (< 0.05) means that the null hypothesis was rejected, and alternate hypothesis was accepted. Therefore, there was at least one useful relationship between an independent variable and the dependent variable. The p-values resulting from the individual t-tests were then assessed. The null and alternate hypotheses for the individual t-tests remained the same as previous regressions. The p-values and whether the null hypothesis can be rejected for each of the independent variables are shown in table 15 below.

Table 15: P-values for the four independent variables.

Independent Variable	P-value	Reject Null Hypothesis
Area (ha)	0.045	Yes
Average Slope (%)	0.053	No
Length/Width Ratio	0.070	No
Boundary Complexity	0.0051	Yes

This shows that boundary complexity had the strongest relationship with landing service area, followed by area. Average slope and length/width ratio were both very close to being significant and allowing the null hypothesis to be rejected, however, they had p-values just greater than 0.05 meaning that we cannot say with absolute confidence that there is a relationship between landing service area and these variables.

The multivariate regression was re-run with the removal of firstly length/width ratio and then average slope, however, both models produced a lower R-squared value and resulted in boundary complexity being the only independent variable that has a significant influence on landing service area.

Based upon the original landing service area regression the two independent variables that have a significant relationship with landing service area are area and boundary complexity. Therefore, line fit plots were developed for these variables against landing service area in order to further display the relationships between the variables. These are shown in figures 20 and 21 below.

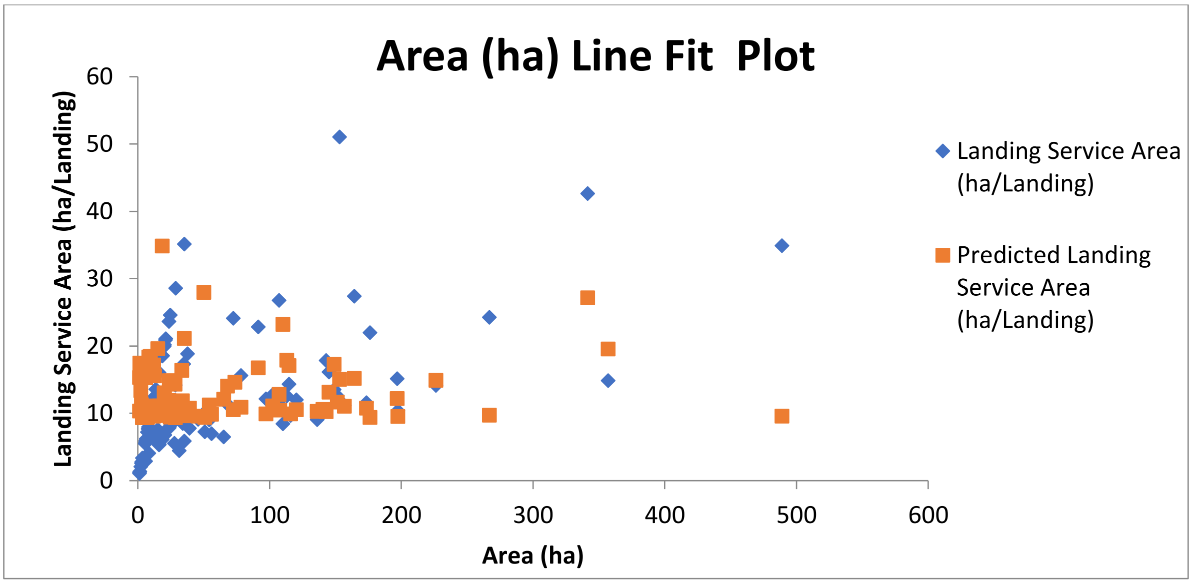

Figure 20: Line fit plot for area against landing service area.

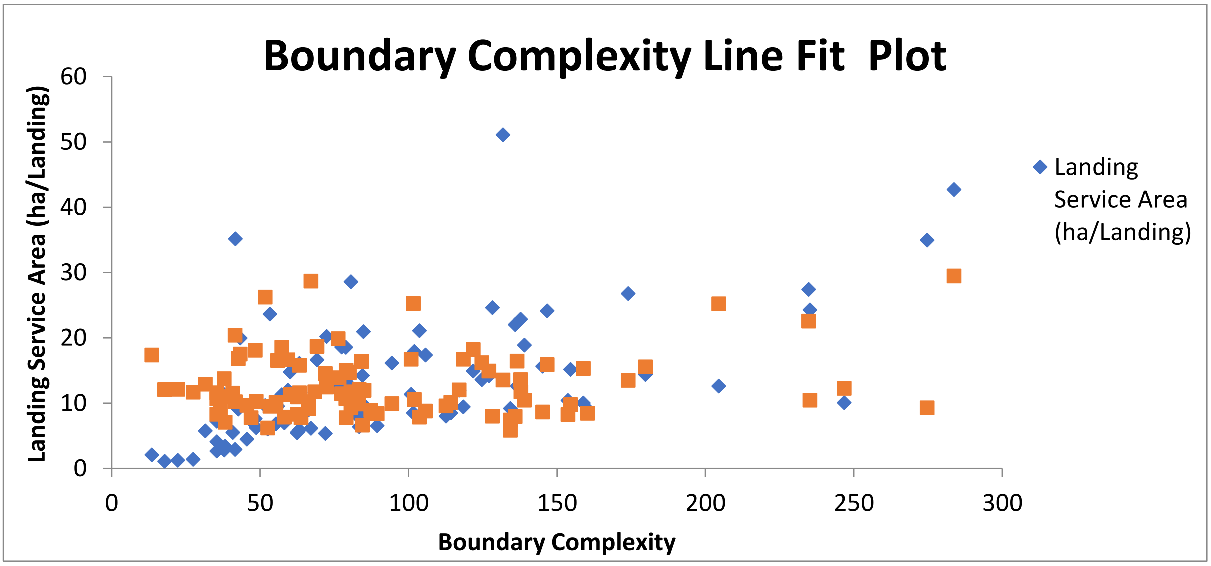

Figure 21: Line fit plot for boundary complexity against landing service area.

Below are the overall regression equations gained from these models, along with the R-squared value for the models, as this represents the strength of the equation. Due to the extremely poor nature of the model that included the 0 m/ha lots it was not included here.

Road Density:

- R-squared = 0.2961

$$Road\ Density\ \left(\frac{m}{ha}\right) = 14.01 + 0.2588 * Avg.Slope + 5.555 * \frac{L}{w} Ratio - 0.06766 * B.Complexity$$

Landing Size:

- R-squared = 0.1876

$$Landing\ Size\ (m^2) = 3913.95 + 2.931 * Area - 20.35 * Avg.Slope - 116.04 * \frac{L}{W} Ratio$$

Average Landing Service Area:

- R-squared = 0.3721

$$ALSA = 10.03 + 0.02633 * Area - 0.07022 * Avg.Slope - 0.6077 * \frac{L}{W} Ratio\ + 0.05707 * B.Complexity$$

5.4 Comparison of Extraction Method means:

Another important independent variable that needed to be assessed for its influence on the infrastructure variables was extraction method. Due to the nature of extraction method not being numerical it was not possible to run a regression analysis involving the data. Therefore, the means of the differing extraction methods were compared using t-tests and box and whisker plots.

Of the 96 woodlots evaluated there were 58 recorded as ground-based extraction, 34 recorded as hauler extraction and 4 recorded as both. The woodlots recorded as both were ignored throughout the following analysis as they were deemed to not add any value to the comparison of extraction methods. The averages of the four infrastructure values for each extraction method are displayed below in table 16.

Table 16: Infrastructure averages for varying extraction methods.

Infrastructure Variable	Ground-based	Hauler
Road Density (m/ha) (All lots)	20.8	32.4
Road Density (m/ha) (No 0m/ha)	27.4	33.4
Landing Area (m^2)	3116.6	2786.9
Average Landing Service Area (ha/landing)	13.1	11.9

14 of the 15 woodlots that were recorded to have no internal roading, and hence 0 m/ha road density, are found in the 58 woodlots that were deemed to have used ground-based extraction. Only 1 of the 15

woodlots are found in the 34 hauler woodlots. This would have significantly lowered the average road density value (all lots) for the ground-based extraction method in comparison to hauler. Due to this, it was decided to only run t-tests and plot box and whisker charts for the road density with no 0 m/ha values included, landing area, and average landing service area.

Box and whisker plots comparing the two extraction methods for the infrastructure variables are shown below in figures 22, 23, and 24. The box and whisker plots show the means, quartiles and ranges of the different extraction methods.

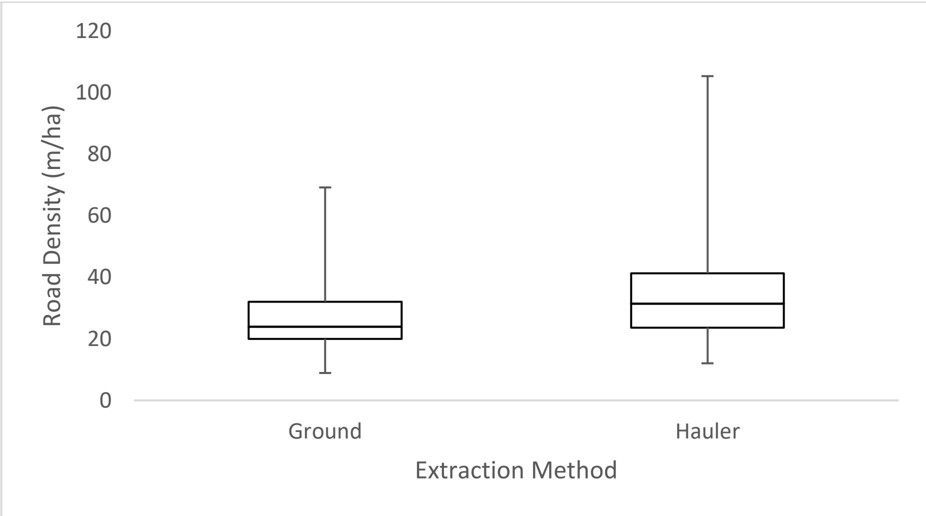

Figure 22: Box and whisker plots comparing road density (m/ha) between extraction methods.

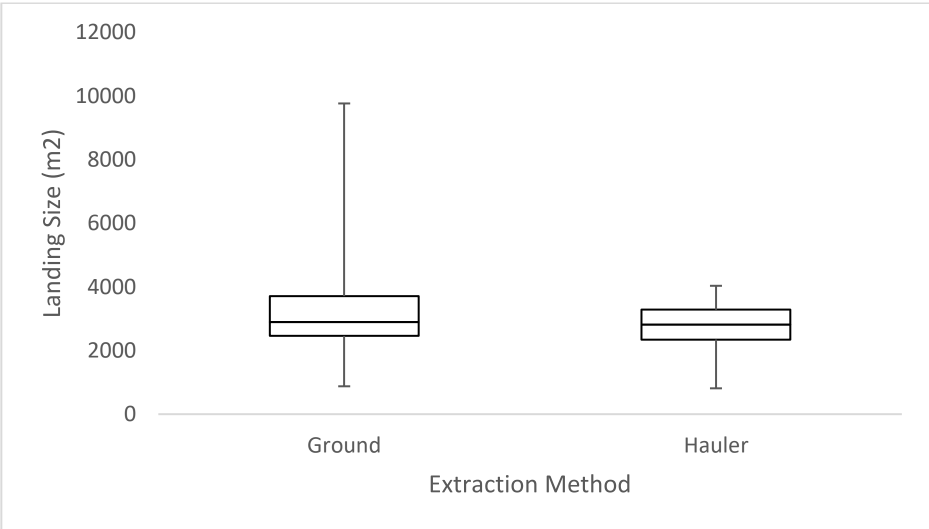

Figure 23: Box and whisker plots comparing landing size (m^2) between extraction methods.

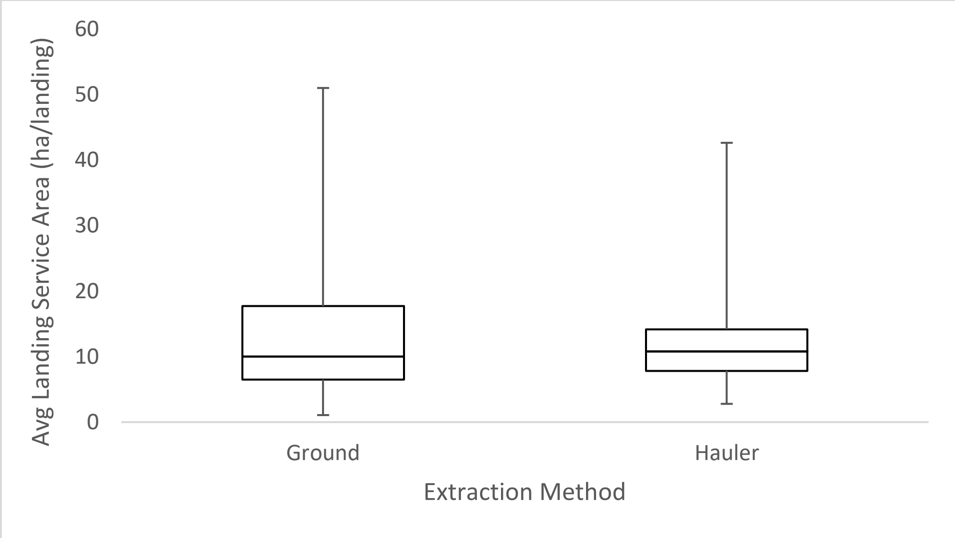

Figure 24: Box and whisker plots comparing landing service area (ha/landing) between extraction methods.

As can be seen in table 16 and figures 22 – 24, the extraction methods produce different averages for the infrastructure variables. To determine whether the differences in means are significant t-tests were conducted. The null and alternate hypotheses for the t-tests were as follows.

Null Hypothesis: H_0: $\mu_1 = \mu_2$ (The means of the two extraction methods are equal).
Alternate Hypothesis: H_1: $\mu_1 \neq \mu_2$ (The means of the two extraction methods are not equal).

The null hypothesis can be rejected, and alternate hypothesis accepted if t-stat > t Critical two-tail.

First a t-test was conducted on roading density, comparing the two means of ground-based and hauler. The results from the t-test assuming equal variance in excel was as follows in table 17.

Table 17: T-test comparing extraction method means for road density.

Extraction methods – Road density t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	*Variable 2*
Mean	27.4	33.4
Variance	185.0	302.8
Observations	44	33
df	75	
t Stat	-1.71	
P(T<=t) two-tail	0.09	
t Critical two-tail	1.99	

Therefore, as seen in table 17; t Stat < t Critical two-tail. This means that we cannot reject the null hypothesis and therefore cannot state that there is a significant difference between the means of road density between extraction methods.

Next a t-test was conducted on landing size. The results from the t-test assuming equal variances are shown below in table 18.

Table 18: T-test comparing extraction method means for landing size.

Extraction methods - landing size t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	*Variable 2*
Mean	3116.6	2786.9
Variance	1837970.0	450128.5
Observations	58	34
df	90	
t Stat	1.32	
P(T<=t) two-tail	0.19	
t Critical two-tail	1.99	

Once again, t Stat < t Critical two-tail. This means that the null hypothesis cannot be rejected once again. Therefore, the landing size means for the two extraction methods are not significantly different.

The third and final t-test conducted was on average landing service area. The results from the third t-test are shown below in table 19.

Table 19: T-test comparing extraction method means for average landing service area.

Extraction methods - Landing service area t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	*Variable 2*
Mean	13.1	11.9
Variance	90.9	56.5
Observations	58	34
df	90	
t Stat	0.59	
P(T<=t) two-tail	0.56	
t Critical two-tail	1.99	

The third t-test produced similar results to the first two. The t stat was found to be less than t Critical two-tail, therefore once again we cannot reject the null hypothesis. This means that the means for the two extraction methods relating to landing service area are not significantly different.

Overall, although differences in the means existed between the two extraction methods for all infrastructure variables, none of the differences could be deemed to be significant.

5.5 Soil Type Comparisons:

The 96 woodlots surveyed throughout this study were found to be located on 26 different soil types. Of these 26 different soil types the highest occurring soil type was Orthic Brown, being recorded 28 times, totalling 29% of woodlots. The second highest occurring type was Firm Brown, which occurred 14 times, this totals 15% of the woodlots surveyed. 9% of woodlots were recorded to be on Orthic Pumice soil. After this, the next highest occurrence is 5 woodlots and then the majority of the other soil types only occur 1 – 3 times. The only comparison of means that was deemed worth making was between Orthic Brown and Firm Brown as they are the only soil types which recorded over 10 woodlots. The mean infrastructure values for Orthic Brown and Firm Brown are shown in table 20 below.

Table 20: Infrastructure values for woodlots on Orthic Brown and Firm Brown soils.

Infrastructure Variable	Orthic Brown	Firm Brown
Road Density (m/ha) (All lots)	24.5	34.1
Road Density (m/ha) (No 0m/ha)	28.5	39.7
Landing Area (m^2)	2995.6	2689.2
Average Landing Service Area (ha/landing)	14.9	13.3

T-tests were run on each variable comparing these two soil types, but no significant differences were found for the infrastructure variables. The outputs from these t-tests between Orthic Brown and Firm Brown soil types for each infrastructure variable are included in Appendix II at the back of this report.

5.6 Regional Comparisons:

Regional comparisons were made for each of the 3 infrastructure values. West Coast was not included in the regional analysis because only 1 sample was taken but the 8 other wood supply regions were included. An ANOVA test was chosen to assess the regional variation in infrastructure variables. The results are shown below in tables 21 – 23. The null and alternate hypotheses for the tests were:

Null Hypothesis: H_0: No difference between the regional means ($\mu_1 = \mu_2 = \ldots = \mu_8$)
Alternate Hypothesis: H_1: At least one of the means are significantly different.

Table 21: Regional ANOVA test for road density.

Anova: Single Factor: Road Density				
SUMMARY				
Groups	*Count*	*Average*	*Variance*	*SD*
Otago	12	26.6	158.6	12.6
Cant	7	25.8	105.5	10.3
Nels/Marl	9	42.8	805.5	28.4
SNI	15	33.0	132.1	11.5
HB	5	35.8	406.8	20.2
EC	5	26.8	42.2	6.5
CNI	16	26.7	138.0	11.7
North	11	24.0	120.0	11.0
ANOVA				
Source of Variation	*df*	*F*	*P-value*	*F crit*
Between Groups	7	1.74	0.11	2.14
Within Groups	72			
Total	79			

Table 22: Regional ANOVA test for landing size.

Anova: Single Factor: Landing Size				
SUMMARY				
Groups	*Count*	*Average*	*Variance*	*SD*
Otago	14	3191.3	1261146	1123.0
Cant	10	2748.7	1474296	1214.2
Nels/Marl	11	3120.6	5305322	2303.3
SNI	18	2661.2	407934	638.7
HB	7	2785.1	559325	747.9
EC	7	2746.6	815752	903.2
CNI	17	3283.2	543511	737.2
North	11	3359.0	775424	880.6
ANOVA				
Source of Variation	*df*	*F*	*P-value*	*F crit*
Between Groups	7	0.75	0.63	2.12
Within Groups	87			
Total	94			

Table 23: Regional ANOVA test for landing service area.

Anova: Single Factor: Landing Service Area				
SUMMARY				
Groups	*Count*	*Average*	*Variance*	*SD*
Otago	14	16.9	132.5	11.5
Cant	10	18.1	105.2	10.3
Nels/Marl	11	15.5	203.5	14.3
SNI	18	8.2	5.4	2.3
HB	7	7.9	23.1	4.8
EC	7	10.7	30.4	5.5
CNI	17	11.7	39.8	6.3
North	11	14.5	23.5	4.9
ANOVA				
Source of Variation	*df*	*F*	*P-value*	*F crit*
Between Groups	7	2.55	0.02	2.12
Within Groups	87			
Total	94			

For the ANOVA tests shown above, the null hypothesis can be rejected if F > Fcrit.

Road Density: F < Fcrit. We cannot reject the null hypothesis. The average road density is not significantly different between regions.

Landing Size: F < Fcrit. We cannot reject the null hypothesis. The average landing size is not significantly different between regions.

Landing Service Area: F > Fcrit. Therefore, we reject the null hypothesis. At least one of the regional means for landing service area is significantly different to another.

6 Discussion:

6.1 Discussion of results:

There were two main objectives for this study. The first was to quantify average values for the amount of infrastructure required to harvest small-scale woodlots between 0 and 500 hectares within New Zealand. The second was to determine which factors influence these averages. Average values were gained for three primary infrastructure values: road density, landing size, and average landing service area. Originally it was planned for the third infrastructure variable assessed to be the number of landings within a woodlot, however, clearly the number of landings increases with increasing size of woodlots. Knowing the area a landing on average services within woodlots in New Zealand has the potential to be much more beneficial to harvest planners and forest owners. This is because the value is in units of hectares/landing; therefore, if a woodlots size (ha) is divided through by the average landing service area value, it should give a rough estimate of how many landings will potentially be required to harvest the given woodlot area. The average values for these infrastructure variables are displayed in table 8 in the results section of this report.

As briefly mentioned in the results section, the road density average varied significantly depending on whether woodlots with no internal roading were included in the averaging. There were 15 woodlots total included in the sample of 96 woodlots that had no internal roading and therefore returned a road density of 0 metres/hectare. In other words, 15.6 % of woodlots surveyed had no internal roading. These woodlots had a landing located on the perimeter of the forest, and all wood was able to be extracted through the use of skid trails or in one case, by a hauler. If the 15 woodlots were included in the road density average, the value was 25.2 m/ha; if they were excluded, the average value returned was 29.9 m/ha. This is a difference in values of 15.6 %. The majority of the woodlots with 0 m/ha road density causing this were very small. 93.3 % (14/15) of the woodlots with no roading were less than 15 hectares in size, and 80% were below 10 hectares in size. Overall, these 15 woodlots had an average area of 6.2 hectares. This result suggests multiple points. It indicates that small-scale forest owners, where profit will not be excessive, are very reluctant to put permanent infrastructure in place. It also likely shows that most small-scale forest owners of this size are only going to be engaging in single rotation forestry. Suppose a forest owner is going to be engaging in multiple rotations. In that case, the cost of permanent infrastructure is spread over multiple rotations making the installation of permanent infrastructure more appealing, rather than the cost being borne for a single rotation. Therefore, if a woodlot is only being operated for a single rotation, and wood can be extracted without the installation of internal roading, then it can be beneficial for saving costs and ensuring profitability. Overall, out of the two values presented for road density the value that does not include the 0 m/ha values is more beneficial. If a woodlot is going to need roading, it gives a more accurate representation of what the average road density may be. This will be more beneficial in estimating how many metres of road may be needed depending on the size of the woodlot.

For the other two infrastructure variables, the following averages were calculated. Average landing size was found to be 3000.1 $m^{2,}$ and the average landing service area was found to be 12.8 ha/landing.

The regional averages shown in table 8, along with the overall national averages just mentioned, need to be used with caution. Based upon Cochran's formula, 96 samples were taken nationally to theoretically achieve a confidence level of 95 ± 10% nationally. It would have been infeasible in the scope of this study to take 96 samples in each region; this means we cannot say that this level of confidence and precision applies to the infrastructure values at a regional level. This lower level of confidence was acknowledged in section 5.6 when the West Coast region was not included in the ANOVA analyses. The other eight wood supply regions were included in the ANOVA tests; however, some of these regions (such as Hawkes Bay and East Coast) still had a relatively low number of samples taken within each region. Therefore, although the ANOVA tests stated that there is not a significant difference between regions for road density and landing size, and there is a significant difference between regions for landing service area, the lower level of confidence needs to be acknowledged. To make any reliable claims, a further study focussing on regional analyses would be necessary. Increasing the number of samples taken within each region would drastically increase our ability to make claims around regional differences. This would also enable t-tests between regions to be completed so that exactly where the difference in means is can be found, as the ANOVA tests do not show where the significant difference lies.

As mentioned in the results section, the level of confidence predicted by Cochran's formula is theoretical, and the true confidence level of the results varied from this slightly. The descriptive statistics function in excel uses the following formula which differs to Cochran's.

$$CI = \overline{x} \pm z * \frac{SD}{\sqrt{n}} \tag{9}$$

Where; CI = Confidence Interval, $\overline{x}$ = the sample mean, z = the z-value for the chosen confidence interval, SD = standard deviation, and n = number of samples.
The experimental 95% confidence intervals are more representative of the actual sample, and therefore the confidence intervals displayed in table 9 should be referred to as the accurate confidence intervals. These differ slightly to the theoretical confidence interval of 95 ± 10 % that Cochran's formula predicted 96 samples would achieve.

The regression analyses and t-tests conducted throughout the results section helped meet the secondary objective of the study, identifying which factors influenced the infrastructure averages. However, they also showed a need for further investigation of other independent variables. There were a total of 6 independent variables assessed to see whether they had a significant influence on the infrastructure variables. Of these 6, 4 were numerical, and 2 were non-numerical. A summary of which of the independent variables had a significant influence on the infrastructure variables is shown below in table 24. The corresponding box is highlighted green if there was a significant relationship and highlighted red if no significant relationship was found. The number within each box corresponds to the rank in terms of which had the most significant relationship, 1 represents the variable with the greatest influence. The road density average with the 0 m/ha values included is not included here as the 0 m/ha woodlots significantly degraded the ability for a significant relationship to be assessed in both the regression analyses and t-tests.

Table 24: Matrix summarising which of the independent variables had a significant relationship with each of the infrastructure variables.

Independent Variable	Road Density	Landing Size	Landing Service Area
Area (ha)		3	2
Average Slope (%)	2	1	
L/W Ratio	1	2	
Boundary Complexity	3		1
Extraction Method			
Soil Type			

Some of the independent variables influenced the dependent variables as expected, some did not. L/W ratio was not predicted to have as much influence as this project has shown, yet it had the strongest relationship with roading density and second strongest relationship with landing size. It likely has the strongest relationship with road density due to some very long woodlots having an entry point at the far end, this then means that roading must be constructed right the way throughout the woodlot. If this is coupled with the woodlot being thin, the result is a small area with a considerable length of road. In other words, a high road density corresponding to a high L/W ratio. This differs to slope which was predicted to have an influence on both road density and landing size and did. When assessing the 96 samples, there were several woodlots located on very flat land that had very large landings in comparison to other woodlots. This is likely due to the fact that the landings were measured post-harvest. It is well known that contractors have a tendency to increase the size of their landing throughout harvest, if space allows. With a flat site it is much easier and safer for a contractor to expand the landing as oppose to a steep site where earth works are likely needed. This could show why slope had the strongest relationship with landing size. Like the strongest variable with road density, it was not predicted that boundary complexity would have the most significant relationship with landing service area.

The average infrastructure values for ground-based vs hauler extraction methods are shown in table 16. When just comparing the averages, we can see that roading density is higher for hauler-based operations, landing area is higher for ground-based operations and average landing service area is higher for ground-based operations also. The increased landing area size for ground-based extraction could relate to the average slope influence on landing size, shown in the regression. However, unexpectedly, the t-tests conducted that compared the means of the infrastructure values for varying extraction methods found the means to not be significantly different. Therefore, although differences between the extraction method averages do exist, they cannot be deemed significantly different. Further investigation into the averages of these extraction methods would be useful in the future.

Problems arose with trying to assess the influence that soil type has on the infrastructure variables. Due to 26 different soil types being returned in the analysis of the 96 woodlots, it made it extremely difficult to compare the means of different soil types. This is because it is unrealistic to assume that the smaller occurring soil types are representative of that soil type as a whole. Soil type correlates very well with location; this was one of the primary reasons for conducting a regional analysis. This again supports the fact that it would be beneficial in the future to expand the study to a wider regional analysis. It would also be beneficial to breakdown the soil types recorded into sands, silts and clays and use these to make a statistical comparison of means.

Although the regression analyses and t-tests showed which of the independent and dependent variables had significant relationships and identified which of these relationships were the strongest, they also highlighted the need for further investigation of more independent variables. At the end of section 5.3 of this report, three equations are listed which are meant to predict the infrastructure variables based upon the independent variable values. However, these were all very weak models. The corresponding R-squared values are summarised in table 25 below.

Table 25: R-squared values corresponding to the regression equations presented in section 5.3.

Model	R-squared value
Road Density (m/ha)	0.2961
Landing Size (m^2)	0.1876
Landing Service Area (ha/landing)	0.3721

As can be seen from the R-squared values, all the models produced from the multivariate regression analyses were relatively week. The highest R-squared value was for landing service area, which still only accounted for 37.2 % of the variability. These low R-squared values show a need to add more independent variables to the study, if more significant variables are added to the regression for each infrastructure variable it would be likely that the R-squared values would increase.

6.2 Comparison of results to other literature:

The results from this study can be compared to other studies mentioned in the review of literature to see if results are consistent. In this study average landing size for woodlots up to 500 hectares in size were found to be 3000.1 m^2. The 2010 study on landing size and characteristics (Visser, Spinelli, & Magagnotti, 2010) found landing size to be 3900 m^2, and the 1987 study (Raymond, 1987) found average landing size to be 2900 m^2. Both of these studies included a range of forest sizes and included plantations larger than 500 hectares in size, therefore a direct comparison cannot be made but the results seem to align relatively well. The 2010 study found the main drivers of landing size to be number of log sorts and production level, whereas the 1987 study found the main drivers to be production, piece length and number of log sorts. None of these variables were assessed for this study as they would have required information around the harvest operations that were conducted in the sampled woodlots. However, the 1987 study also found extraction method not to have an influence on landing size, the same as this study. The influences found in the 1987 and 2010 studies could explain why for landing size the regression model produced the lowest R-squared value of all models at only 0.1876. Including the significant variables found in these studies would likely greatly raise this R-squared value. Similar to our study, the 1987 study also found that a lot of the variation of landing size, could not be explained by the variables assessed in this study. An interesting finding from the 1987 study was that soil type did have an influence, which this study did not. This is likely due to the comparison being done through a regional analysis that was based upon the prominent soils in those regions, rather than the specific soils under the sites sampled. This again highlights the need for further regional analysis.

Literature in New Zealand that outlines values for average road density was hard to find. However, some studies outlined factors that influence road density. A road engineering manual produced in Ireland (Ryan, Phillips, Ramsay, & Dempsey, 2004) stated that road density relates to the planned harvesting and extraction methods used. However, this study did not find that extraction method influenced the road density of woodlots.

6.3 Limitations to the study:

There were several limitations to this study. One of the main limitations was the delay in mapping of woodlots nationally that lead to Auckland not being included in the study. Auckland is part of the Northland wood supply region and samples were still able to be taken within that region, as well as some which were just outside Auckland's regional boundaries. Therefore, it was deemed to not have a significant influence on the study. However, it is still a limitation that it was not possible for samples to be randomly taken from within that area.

Another limitation was woodlot clarity issues that occurred in Google Earth. Clarity issues that resulted in a woodlot not being able to be assessed for its infrastructure include; year gaps in historical imagery (i.e. the woodlot was in awaiting Woodstock status in 2016, yet historical imagery could only be accessed for 2014 and 2018), cloud cover, shadowing from surrounding topography, dark blue/black image quality, and blurred landing boundaries due to slash or erosion. If any of these clarity issues were encountered when trying to collect data from a woodlot it meant that the woodlot needed to be resampled. In some cases (for example with cloud cover or poor blue/black image quality) this meant that all woodlots within a certain area had poor clarity.

For this study woodlot size was restricted to 500 hectares, although some literature defines a woodlot as a forest up to 1000 hectares. This was done for several reasons; the number of woodlots in the 500 – 999 ha range are very limited, and even fewer would have been harvested recently. Also, when determining extraction method, the larger the woodlot area the greater possibility a range of equipment will be used. Therefore, restricting the size to 500 hectares was deemed to make extraction method determination easier.

A limitation to this study was also that landing boundary could be very difficult to determine in certain cases. If there was excessive slash stored on the edge of a landing or erosion had occurred on steep sites it meant the boundary was blurred and determining the exact point the landing ended became difficult. In these cases professional judgement and historical imagery was used to help assist in determining the boundary, but some level of human error should be expected in these cases.

As discussed earlier, the limited number of samples taken in each region was another limitation of the study which resulted in the regional comparisons made being limited in their guaranteed accuracy.

7 Conclusion:

The main objectives of this study were to first quantify average values for the amount of infrastructure required to harvest small-scale woodlots in New Zealand, and secondly, evaluate what variables are driving these averages.

The primary objective was met and the infrastructure averages gained were as follows; road density (0 m/ha samples included) = 25.2 metres/hectare, road density (0 m/ha samples excluded) = 29.9 metres/hectare, landing size = 3000.1 m^2, and landing service area = 12.8 hectares/landing. Out of the two road density values reported, the second value is more beneficial to the forestry industry as it should be determined early on in the planning stage if roading will be needed within a woodlot. In the case that it does, it does not make sense to use an average value that includes 15 samples that recorded a road density of 0 metres/hectare, in turn lowering the average.

The secondary objective was also met; however, only some of the variables influencing the average values stated above were found. The 6 variables assessed for their influence were; area (ha), average slope (%), length/width ratio, boundary complexity, extraction method, and soil type. From the most substantial relationship to the weakest, the following relationships were found. Road density has a significant relationship with; length/width ratio, average slope (%), and boundary complexity. Landing size has a significant relationship with; average slope (%), length/width ratio, and area (ha). Finally, landing service area has a significant relationship with; boundary complexity and area (ha).

Overall, the results gained throughout this study can be deemed to be accurate and reliable. The average values for infrastructure gained in this study have the potential to be beneficial to industry in terms of assisting in costing estimates, assisting in harvest planning, and supplementing further research. Through finding which of the assessed variables influence the averages, the forestry industry could be further benefited by using these relationships to assist in cost reduction, which is becoming increasingly important as the woodlots planted in the 1990s planting boom mature.

7.1 Further research:

A further investigation of soil types could be beneficial. 26 different soil types were recorded during this study, making comparisons difficult. By breaking down these soil types into silts, sands and clays it would make comparisons easier and be a more beneficial comparison as silts, sands and clays differ significantly in terms of their workability and structural integrity. By breaking down soils into these classes, they could also be linked with a regional study in terms of the prominent soil type in certain areas.

A regional analysis is another aspect that would be good to be explored. As mentioned throughout section 6, many more samples need to be taken in each wood supply region to enable a reliable full-scale comparison of infrastructure requirements between regions. This could then be linked with likely costs by region. If a further regional analysis were conducted, it would also be useful to evaluate woodlots within the Auckland region; this would allow clarity around whether the exclusion of the Auckland region impacted this study.

In order to increase the strength of the regression equations gained for predicting infrastructure values, more variables should be assessed for their influence. A focus should be put on the variables listed by other literature that could not be assessed in the scope of this project, especially if the study is striving to gain equations that account for a high level of variability in the dependent variables.

Finally, if a similar study were completed on large commercial scale forestry plantations, it would be interesting to make comparisons between the values gained for woodlots, and the values gained for commercial plantations.

References

Brown, K., & Visser, R. (2018). Adoption of emergent technology for forest road management in New Zealand . *NZ Journal of Forestry, Vol. 63, No. 3*, 23-29.

Farah, A., & Algarni, D. (2014). Positional Accuracy Assessment of Google Earth in Riyadh. *Artificial Sattelites, Vol. 49(2)*, 101-106.

Forest Owners Association (FOA). (2019). *Facts & Figures 2018/19.* Wellington: Forest Owners Association.

Ghaffarian, M. R., & Sobhani, H. (2008). Optimum road spacing of ground based skidding operations in Nowshahr, Iran. *Caspian Journal of Environmental Sciences*, 105-112.

Harrington, S., Teitelman, J., Rummel, E., Morse, B., Chen, P., Eisentraut, D., & McDonough, D. (2017). Validating Google Earth Pro as a Scientific Utility for Use in Accident Reconstruction. *SAE International Journal of Transport Safety, 5(2)*, 135-166.

Indufor Asia Pacific. (2016). *Wood Availability Forecasts – New Zealand 2014 - 2050.* Wellington: Ministry for Primary Industries (MPI).

Laurie, A. (2020). *Monthly Market Report.* Christchurch: Laurie Forestry Limited.

Manaaki Whenua - Landcare Research. (2020, May). *Soils Map Viewer*. Retrieved from Manaaki Whenua - Landcare Research: https://soils-maps.landcareresearch.co.nz/

Manley, B., Morgenroth, J., Visser, R., & BForSc Students, 2. a. (2017). Quantifying the small-scale owners' estate in Canterbury, Otago and Southland . *NZ Journal of Forestry, Vol. 62, No. 3*, 24 - 32.

Manley, B., Morgenroth, J., Xu, C., & BForSc Students, 2. a. (2020). Quantifying the area of the small-scale owners' forest estate in the East Coast, Hawke's Bay and Southern North Island. *NZ Journal of Forestry, Vol. 64, No. 4*, 18-26.

Ministry for Primary Industries (MPI). (2019). *National Exotic Forest Description.* Wellington: Ministry for Primary Industries (MPI).

Ministry for Primary Industries (MPI). (March, 2020). *Indicative New Zealand Radiata Pine Log Prices.* Wellington: Ministry for Primary Industries (MPI).

Ministry of Primary Industries (MPI). (2020). *Situation and Outlook for Primary Industries.* Wellington: Ministry for Primary Industries (MPI).

New Zealand Forest Owners Association Incorporated. (2012). *NZ Forest Road Engineering Manual.* Wellington: New Zealand Forest Owners Association Incorporated.

Petherick, K. (2014). *Assessing the Level of Unproductive Area in Production Forestry Sites in the South Island of New Zealand.* Christchurch: Universtiy of Canterbury.

Rafiei, A. A., Lotfalian, M., Hosseini, S. A., & Parsakhoo, A. (2009). Determining the Optimum Road Density for Ground Skidding System in Dalak Kheyl Forest-Hycranian Zone. *World Applied Sciences Journal 7 (3)*, 263-270.

Raymond, K. A. (1987). *Factors Influencing Landing Size.* Rotorua: Logging Industry Research Association.

Ryan, T., Phillips, H., Ramsay, J., & Dempsey, J. (2004). *Forest Road Manual - Guidelines for the design, construction and management of forest roads.* Dublin: COFORD, National Council for Forest Research and Development.

Visser, R., & Murphy, G. (2019). *How to market and harvest your forest woodlot for profit. .* Rotorua, Christchurch: Forest Growers Research Ltd.

Visser, R., Spinelli, R., & Magagnotti, N. (2010). *Landing size and landing layout in whole-tree harvesting operations in New Zealand.* Padova, Italy: FORMEC.

Appendix I:

Multivariate regression 1:

Road Density - All Lots Regression

Regression Statistics	
Multiple R	0.288120133
R Square	0.083013211
Adjusted R Square	0.042706099
Standard Error	17.30111098
Observations	96

ANOVA

	df	SS	MS	F	Significance F
Regression	4	2465.888923	616.4722	2.059517728	0.09261568
Residual	91	27238.88813	299.3284		
Total	95	29704.77706			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	13.68704719	5.852237078	2.338772	0.021537048	2.062297158	25.31179723	2.062297158	25.31179723
Area (ha)	0.015417354	0.031591734	0.488019	0.626710472	-0.047335745	0.078170453	-0.047335745	0.078170453
Average Slope (%)	0.228857551	0.087452907	2.616923	0.010389389	0.055143094	0.402572008	0.055143094	0.402572008
L/W Ratio	0.333686665	0.810701382	0.411602	0.681598686	-1.276672039	1.944045369	-1.276672039	1.944045369
Boundary Complexity	0.003051936	0.048540848	0.062874	0.950005089	-0.093368495	0.099472366	-0.093368495	0.099472366

Multivariate regression 2:

Road Density - 0 m/ha lots NOT included

Regression Statistics	
Multiple R	0.544229554
R Square	0.296185807
Adjusted R Square	0.259142955
Standard Error	13.06013953
Observations	81

ANOVA

	df	SS	MS	F	Significance F
Regression	4	5455.259937	1363.814984	7.995761371	1.95758E-05
Residual	76	12963.11058	170.5672444		
Total	80	18418.37052			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	13.84074348	5.483113679	2.524248864	0.013680413	2.920176146	24.76131081	2.920176146	24.76131081
Area (ha)	-0.002440414	0.024174982	-0.100947916	0.919857647	-0.050589063	0.045708235	-0.050589063	0.045708235
Average Slope (%)	0.259602149	0.07201917	3.604625697	0.000556146	0.116163542	0.403040757	0.116163542	0.403040757
L/W Ratio	5.576267055	1.355681002	4.113258978	9.78864E-05	2.876194335	8.276339775	2.876194335	8.276339775
Boundary Complexity	-0.064885553	0.038254862	-1.696138734	0.093952453	-0.141076714	0.011305608	-0.141076714	0.011305608

Multivariate regression 2 repeat excluding area:

Re-run: Road Density - 0 m/ha lots NOT included (Area Excluded)

Regression Statistics	
Multiple R	0.544142845
R Square	0.296091436
Adjusted R Square	0.268666427
Standard Error	12.97592613
Observations	81

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	5453.521775	1817.840592	10.79640252	5.33335E-06
Residual	77	12964.84874	168.374659		
Total	80	18418.37052			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	14.0118009	5.18108162	2.704416167	0.008418683	3.694949348	24.32865244	3.694949348	24.32865244
Average Slope (%)	0.258843068	0.071163685	3.637291503	0.000495962	0.117138067	0.400548069	0.117138067	0.400548069
L/W Ratio	5.55506757	1.330680621	4.1746062	7.78421E-05	2.905343935	8.204791206	2.905343935	8.204791206
Boundary Complexity	-0.067664363	0.026393532	-2.563672153	0.01230502	-0.120220599	-0.015108126	-0.120220599	-0.015108126

Multivariate regression 3:

Landing Size Regression

Regression Statistics	
Multiple R	0.439972521
R Square	0.193575819
Adjusted R Square	0.158128602
Standard Error	1039.885616
Observations	96

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	23621098.88	5905275	5.460959612	0.000549805
Residual	91	98403950.56	1081362		
Total	95	122025049.4			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	3750.237321	351.7495013	10.66167	1.06799E-17	3051.530148	4448.944494	3051.530148	4448.944494
Area (ha)	1.772018452	1.898825446	0.933218	0.353176881	-1.999765034	5.543801937	-1.999765034	5.543801937
Average Slope (%)	-20.3095445	5.256368824	-3.8638	0.00020894	-30.75067549	-9.86841359	-30.75067549	-9.86841359
L/W Ratio	-108.758649	48.72731622	-2.23199	0.02807071	-205.549477	-11.9678215	-205.549477	-11.9678215
Boundary Complexity	2.386271684	2.917554206	0.817901	0.415549952	-3.409091339	8.181634706	-3.409091339	8.181634706

Multivariate regression 3 repeat excluding boundary complexity:

Re-run: Landing size regression (excluding boundary complexity)

Regression Statistics	
Multiple R	0.433183105
R Square	0.187647602
Adjusted R Square	0.16115785
Standard Error	1038.013055
Observations	96

ANOVA

	df	SS	MS	F	Significance F
Regression	3	22897707.97	7632569	7.083780996	0.000246066
Residual	92	99127341.47	1077471		
Total	95	122025049.4			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	3913.952354	288.7226154	13.5561	1.18564E-23	3340.524318	4487.38039	3340.524318	4487.38039
Area (ha)	2.931012097	1.261667912	2.323125	0.022376716	0.425230721	5.436793472	0.425230721	5.436793472
Average Slope (%)	-20.3543738	5.24661823	-3.87952	0.000196455	-30.77461053	-9.93413717	-30.77461053	-9.93413717
L/W Ratio	-116.044293	47.81993061	-2.42669	0.017186111	-211.0188028	-21.0697831	-211.0188028	-21.0697831

Multivariate regression 4:

Landing Service Area Regression

Regression Statistics	
Multiple R	0.610037839
R Square	0.372146164
Adjusted R Square	0.344548194
Standard Error	7.084319563
Observations	96

ANOVA

	df	SS	MS	F	Significance F
Regression	4	2707.027543	676.7568856	13.48454809	1.13817E-08
Residual	91	4567.070114	50.18758367		
Total	95	7274.097657			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	10.0275457	2.396326899	4.184548318	6.57859E-05	5.267536688	14.7875547	5.267536688	14.7875547
Area (ha)	0.026332966	0.012935929	2.035645587	0.044693896	0.000637333	0.0520286	0.000637333	0.0520286
Average Slope (%)	-0.070220775	0.035809512	-1.96095314	0.052942116	-0.141351972	0.000910422	-0.141351972	0.000910422
L/W Ratio	-0.607669808	0.331959472	-1.83055421	0.070440312	-1.267066519	0.051726902	-1.267066519	0.051726902
Boundary Complexity	0.057066112	0.019876115	2.871089821	0.005087693	0.017584651	0.096547573	0.017584651	0.096547573

Appendix II:

T-test comparing road density in woodlots on Orthic Brown and Firm Brown soil types:

t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	*Variable 2*
Mean	28.54168	39.72867604
Variance	122.34	646.7063292
Observations	24	12
Pooled Variance	291.988	
Hypothesized Mean Difference	0	
df	34	
t Stat	-1.85172	
P(T<=t) one-tail	0.036383	
t Critical one-tail	1.690924	
P(T<=t) two-tail	0.072766	
t Critical two-tail	2.032245	

T-test comparing landing size in woodlots on Orthic Brown and Firm Brown soil types:

t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	*Variable 2*
Mean	2995.565	2689.203176
Variance	2311545	1337396.092
Observations	28	14
Pooled Variance	1994947	
Hypothesized Mean Difference	0	
df	40	
t Stat	0.662654	
P(T<=t) one-tail	0.255676	
t Critical one-tail	1.683851	
P(T<=t) two-tail	0.511352	
t Critical two-tail	2.021075	

T-test comparing average landing service area in woodlots on Orthic Brown and Firm Brown soil types:

t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	*Variable 2*
Mean	14.92487	13.28754
Variance	141.447	120.8298
Observations	28	14
Pooled Variance	134.7464	
Hypothesized Mean Difference	0	
df	40	
t Stat	0.43092	
P(T<=t) one-tail	0.33442	
t Critical one-tail	1.683851	
P(T<=t) two-tail	0.66884	
t Critical two-tail	2.021075	