Master Thesis in Geographical Information Science nr 166

Tree species impact on Forest Fire Spread Susceptibility in Sweden

Sara Sharon Jones

2023
Department of
Physical Geography and Ecosystem Science
Centre for Geographical Information Systems
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Sara Sharon Jones (2023). Tree species impact on Forest Fire Spread Susceptibility in Sweden**.**
Master degree thesis, 30 credits in Master in Geographical Information Science
Department of Physical Geography and Ecosystem Science, Lund University

Tree species impact on Forest Fire Spread Susceptibility in Sweden

Sara Sharon Jones

Master thesis, 30 credits, in Geographical Information Sciences

Supervisors

Veiko Lehsten

Department of Physical Geography and Ecosystem Science Lund University

Igor Drobyshev

Southern Swedish Forest Research Centre, Swedish University of Agricultural Sciences

Acknowledgements

First I would like to thank Igor Drobyshev & Veiko Lehsten for their valuable feedback, guidance and supervision. Thank you to the Southern Swedish Forest Research Centre and the Oasis Project for the opportunity to contribute to fire research in Sweden.

I would like to extend my thanks to the following people:

To my colleagues; Maksym Matsala for his quick mind, problem solving with me and his constant support; Guilherme Pinto for his feedback, advice and kindness; Benjamin Forsmark for always listening and sharing his invaluable knowledge.

To my fellow students Kathrin Bögelsack, Bernat Dorado-Guerrero and Gabriel Romeo Ferriols Pavico, my thanks for forming the community and support system to undertake an entirely distance learning MSc., without whose support this MSc. would not have been possible.

To Dr. Thomas Groen for his outstanding course on Spatial Biological Modelling, through which I learned R and modelling.

To Martin Villet, for believing in me all those years ago.

Finally, thank you to my partner Andreas Ridderstedt for his infinite patience and compassion.

Abstract -EN

Climate change is expected to result in longer fire seasons and increased fire risk. In Sweden, where homogenous production forest makes up the majority of the landscape, there is a need for better understanding of forest fire susceptibility in order to mitigate risk. In the context of this thesis forest fire susceptibility is the likelihood of forest burning after ignition and initial fire propagation has taken place. Forest fire susceptibility contributes directly to the risk posed to the environment, society and economy by fire.It is well established that the main factors governing fire risk are climate and available fuel, of these, only fuel can be feasibly manipulated and managed. It is therefore important to investigate forest fire susceptibility within the context of forest parameters.

This study assessed the impact of several forest parameters on forest fire susceptibility, using hypothesis testing and Generalised Linear Models (logistic regression models). Data consisted of delineated fires from 2018, 2019 and 2020 and the SLU Forest Map rasters.

Hypothesis testing identified significant differences between the burned and unburned classes of tree age (yrs), tree height (m) and standing biomass volume of scots pine ($m^3sk\ ha^{-1}$), birch (*Betula spp.)* ($m^3sk\ ha^{-1}$) and total deciduous ($m^3sk\ ha^{-1}$). After correction for multiple testing only tree age and standing biomass volume of scots pine retained significance, with birch and total deciduous showing marginal significance.

Two logistic regression models were trained with 5-fold cross validation, while also holding back 20% of the dataset for accuracy testing. Model 1 was trained with age, and standing biomass volume of total deciduous, pine and spruce. Model 2 was trained with and standing biomass volume of pine and spruce. Accuracies for both models were low, with AUC values of 0.556 for model 1 and 0.551 for model 2. In model 1 total deciduous standing biomass volume and age had no significant contribution to forest fire susceptibility, standing biomass volume of pine and spruce had at best a marginal contribution to forest fire susceptibility (p values of 0.078 and 0.060 respectively), where pine increased and spruce decreased forest fire susceptibility. After adjusting p values for multiple testing neither pine nor spruce contributed significantly (adjusted p values of 0.312 and 0.240 respectively). In model 2 spruce standing biomass volume was found to marginally decrease forest fire susceptibility (p value of 0.034 and adjusted p value of 0.067) and pine standing biomass volume had no significance (adjusted p value of 0.121) on forest fire susceptibility.

Hypothesis testing highlighted different variables of importance than those of the logistic regression models. Spruce standing biomass volume was the only variable with any significant contribution to forest fire susceptibility after adjusting p-values for multiple testing (model 2), which had no significant difference between burned and unburned classes in hypothesis testing.

More research is needed with a larger dataset and improved sampling methods to fully understand the relationships between forest parameters and forest fire susceptibility in Sweden.

Keywords: Geography, GIS, Fire, Forest Fire Susceptibility, Sweden, logistic regression modelling

Advisors: Igor Drobyshev & Veiko Lehsten
Master degree project, 30 credits Geographical Information Sciences, 2023
Department of Physical Geography and Ecosystem Science, Lund University
Thesis nr 166

Abstract -SV

Klimatförändringar förväntas leda till längre brandperioder och ökad brandrisk. I Sverige, där homogen produktionsskog utgör största delen av landskapet, finns ett behov av bättre förståelse för skogens mottaglighet för bränder för att kunna mildra risker. I denna avhandling avser skogsbrandmottaglighet sannolikheten för att skogen brinner efter att antändning har ägt rum och inledande brandspridning har skett. Skogsbrandmottaglighet bidrar direkt till den risk som bränder utgör för miljö, samhälle och ekonomi. Det är väl fastställt att de huvudsakliga faktorerna som styr brandrisk är klimat och tillgängligt bränsle; av dessa kan endast bränsle manipuleras och hanteras. Det är därför viktigt att undersöka skogsbrandmottaglighet i sammanhanget av skogsparametrar.

I denna studie utvärderades effekten av flera skogsparametrar på skogsbrandmottaglighet med hjälp av hypotesprövning och generaliserade linjära modeller (logistiska regressionsmodeller). Data bestod av avgränsade bränder från 2018, 2019 och 2020 samt SLU Skogskartans raster.

Utförd hypotesprövning identifierade signifikanta skillnader mellan brända och obrända klasser av trädålder (år), trädhöjd (m) och stående biomassa av tall ($m^3sk\ ha^{-1}$), björk (Betula spp.) ($m^3sk\ ha^{-1}$) och total lövskog ($m^3sk\ ha^{-1}$). Efter korrigering för multipel testning behöll endast trädålder och stående biomassa av tall signifikans, medan björk och total lövskog visade marginell signifikans.

Två logistiska regressionsmodeller tränades med 5-faldig korsvalidering, med 20% av datamängden undanhållen för noggrannhetstestning. Modell 1 tränades med ålder samt stående biomassa av total lövskog, tall och gran. Modell 2 tränades med stående biomassa av tall och gran. Noggrannheten för båda modellerna var låg, med AUC-värden på 0,556 för modell 1 och 0,551 för modell 2. I modell 1 hade total lövskogs stående biomassa och ålder inget signifikant bidrag till skogsbrandmottaglighet, stående biomassa av tall och gran hade vid bästa marginal en bidragande betydelse till skogsbrandmottaglighet (p-värden på 0,078 respektive 0,060), där tall ökade och gran minskade skogsbrandmottagligheten. Efter justering av p-värden för multipel testning bidrog varken tall eller gran signifikant (justerade p-värden på 0,312 respektive 0,240). I modell 2 visade sig stående biomassa av gran marginellt minska skogsbrandomottagligheten (p-värde på 0,034 och justerat p-värde på 0,067), och stående biomassa av tall hade ingen signifikans (justerat p-värde på 0,121) för skogsbrandmottaglighet.

Hypotesprövning framhöll olika variabler av betydelse jämfört med de som framkom i de logistiska regressionsmodellerna. Stående biomassa av gran var den enda variabeln med något signifikant bidrag till skogsbrandomottaglighet efter justering av p-värden för multipel testning (modell 2), vilken inte hade någon signifikant skillnad mellan brända och obrända klasser i hypotesprövning.

Mer forskning krävs med en större datamängd och förbättrade urvalsmetoder för att fullständigt förstå sambanden mellan skogsparametrar och skogsbrandmottaglighet i Sverige.

Nyckelord: Geografi, GIS, Brand, Skogsbrandomottaglighet, Sverige, logistisk regressionsmodellering

Handledare: Igor Drobyshev & Veiko Lehsten
Examensarbete för magisterexamen 30 hp i geografisk informationsvetenskap, 2022
Institutionen för naturgeografi och ekosystemvetenskap, Lunds universitet Examensarbete nr 166

Contents

List of Figures

List of Tables

1. Introduction

The importance of Fire

Prior to the Anthropocene, wildfire activity was controlled by climate and modulated by topography, fuels, and ignition patterns (Pinto et al. 2020a). Today wildfire activity is heavily influenced by human activity, via fire suppression and anthropogenic climate change (Drobyshev et al. 2012; Molinari et al. 2020; Pinto et al. 2020b). In Sweden fire plays an important role in forest ecosystems (Johnson et al. 2003; Hörnberg et al. 2004), driven by climate variability and over the last 400 years – by both climate and humans (Niklasson and Granström 2000). Fire activity has declined since the 1800's due to the rise in timber production and subsequent introduction of fire suppression (Johnson et al. 2003; Hörnberg et al. 2004). Currently, an average of 3600 ha is burnt annually in Sweden, in around 5000 fires (Sjöström and Granström 2020).

The fire season is the period when fires occur. It commonly starts when the daily temperature reaches 9^0C for 3 consecutive days and ends when the temperature drops below 2^0C (Flannigan et al. 2013). In Sweden the fire season typically starts in April and ends in October (Pinto et al. 2020b).

Climate change is likely to result in longer fire seasons and subsequent increases in fire risks to various forest ecosystem functions (Pinto et al. 2020a). As a result, forest fires pose an increasing economic risk to the forest industry.

Forest Fire in Sweden

Forestry is an important part of Sweden's economy, with production forest covering 58% of the surface area of the country (Nilsson et al. 2021) the importance of understanding forest fire susceptibility in Sweden is clear. Conifers make up 82% of the production forest area, Scots pine *Pinus sylvestris* (39.8%) and Norway spruce *Picea abies* (27.7%) are the predominant forestry species (Nilsson et al. 2021). Non-production forests make up just under 1% of the surface (Roberge et al. 2020).

Forest flammability is driven by many factors, such as vegetative moisture content, fuel load, fuel type, fuel moisture content, surface area to volume ratio of fuels, groundcover, understory and canopy level species, crown height and planting density (Van Wagtendonk et al. 1996; Päätalo 1998; Hély et al. 2001; Scott and Burgan 2005; Hély et al. 2010; Xanthopoulos et al. 2012; Chuvieco et al. 2014). Ignition is needed to start a fire, and propagation will occur only if there is a suitable fuel source.

Conifer forests are more fire-prone than deciduous forests, and monodominant coniferous stands, which are predominant in Sweden, are more fire-prone that mixed stands (Päätalo 1998; Terrier et al. 2013; Pinto et al. 2020b). Variations in leaf moisture content are primarily the reason for differences in tree flammability, the water content of broadleaves is 150% of their dry weight, in comparison to only 100% of conifer needle foliage (Päätalo 1998). Conifer forests also have slower leaf litter decomposition, resulting in a build-up of surface fuel, and the understory is typically characterized by flammable lichens and mosses (Päätalo 1998). Stand age plays a role in flammability, with young trees of all species being more flammable than mature trees due to increased stand density and proximity of the canopy to the surface fuels (Päätalo 1998).

Over-mature stands, where natural tree death occurs more frequently, are also more susceptible to fire (Päätalo 1998).

The differences in flammability are strongly controlled by average fuel moisture content and are particularly pronounced under the conditions of moderate drought. It is thought that extreme drought homogenizes fuel conditions across stand types, making both deciduous and coniferous stands conducive to burning (personal communication with I. Drobyshev, 2021).

Norway spruce are both flammable and highly susceptible to damage during a fire, the species is well adapted to moist sites and lower light conditions, as its crown extends down low enough to provide fuel connectivity to fuels on the forest floor. This increases the risk of surface fires spreading into the crowns of trees (Päätalo 1998).

Scots pine are resilient to fire damage, with a high survivability (Päätalo 1998). They are less prone to canopy fires due to a high shallow crown, self-pruning reducing ladder fuels, thick bark and heat insulating properties reducing likelihood of surface fire spreading to the canopy (Päätalo 1998). Scots pine are planted on xeric sites, where they are well adapted, these drier sites have an increased likelihood of surface fires and Scots pine burns readily under dry conditions (Päätalo 1998).

Coniferous forests feature the highest rate of spread, fire intensity and burned area, as compared to mixed and deciduous forests in North America (Hély et al. 2000; Terrier et al. 2013). In Sweden this comparison has not yet been as thoroughly investigated.

In Sweden the highest fire spread rates are seen in mixed species conifer stands of Scots pine and Norway spruce, whereas deciduous and mixed stands have lower susceptibility to fires and lower fire spread rates (Päätalo 1998). The more deciduous trees a stand contains, the lower the rate of fire spread (Päätalo 1998). Mitigation of fire risk could be achieved by planting deciduous trees in traditionally conifer dominated forest production (Terrier et al. 2013). The susceptibility of forest type to fire has not yet been quantified.

1.1. Research Objective

With climate change and an increase in extreme weather events, fires are becoming more frequent and extensive in Sweden. In Sweden, the majority of forest fires occur in production forests. This poses a risk to the forestry industry, which relies heavily on monoculture stands of Scots Pine and Norway spruce for timber production, two conifer species believed to be high risk to forest fires.

1.2. Research Objectives & Questions and Hypotheses

The project aims to investigate the relationship between forest fire spread susceptibility and tree species as well as other forest parameters, in Swedish fires above 10ha with data collected from burned areas and unburned fire edges. This was carried out using paired hypothesis tests to establish whether each parameter varies significantly between burned and unburned classes, and logistic regression modelling was used to identify which parameters contribute or reduce burn likelihood. The research questions were defined as follows:

Q1. Is there a significant difference between forest parameters in burned and unburned areas?

I hypothesise that there is a significant difference between the values of the parameters in burned and unburned areas for all parameters.

Q2.1. Which forest parameters significantly contribute to burn likelihood in the logistic regression model?

I hypothesise that increases in pine volume, spruce volume, above-ground biomass and total tree volume will significantly increase burn likelihood.

Q2.2. Which forest parameters significantly reduce burn likelihood in the logistic regression model?

I hypothesise that increases in oak, beech, birch, other deciduous and total deciduous volume will significantly decrease burn likelihood.

2. Methods

2.1 Study Area

Sweden is a Northern European country located between the latitudes of 55^0 and 70^0N, with a land area of 40.7 million ha. The South experiences a milder oceanic climate due to the warm North Atlantic Current (Roberge et al. 2020) and the North Atlantic westerlies (Wastenson et al. 1995), whereas the North has a continental climate. While summer temperatures of the North and South do not vary much, winter temperatures in the North are much colder, with a mean January temperature of -14^0C in comparison to 0^0C in the South (Wastenson et al. 1995). There is a strong East-west precipitation gradient, the east receives an annual precipitation of 500-6000mm, and the West around 1100mm westerlies (Wastenson et al. 1995).

Sweden's natural vegetation spans six biogeographical zones from alpine to nemoral (Ahti et al. 1968; Wastenson et al. 1995). Around 69% of Sweden is forested (27.9 million ha), of which the majority is production forest, accounting for around 58% of Sweden's surface area (23.5 million ha), (Roberge et al. 2020). Non-production forests make up just under 1% of the surface (Roberge et al. 2020).Conifers make up 82% of the production forest area, Scots pine *Pinus sylvestris* (39.8%) and Norway spruce *Picea abies* (27.7%) are the predominant forestry species (Nilsson et al. 2021).

2.2. Software

ESRI ArcGIS Desktop version 10.5.1 was used to manage geospatial data (ESRI 2017) and R version 3.6.3 (The R Foundation for Statistical computing 2020) was used to carry out the analyses and modelling.

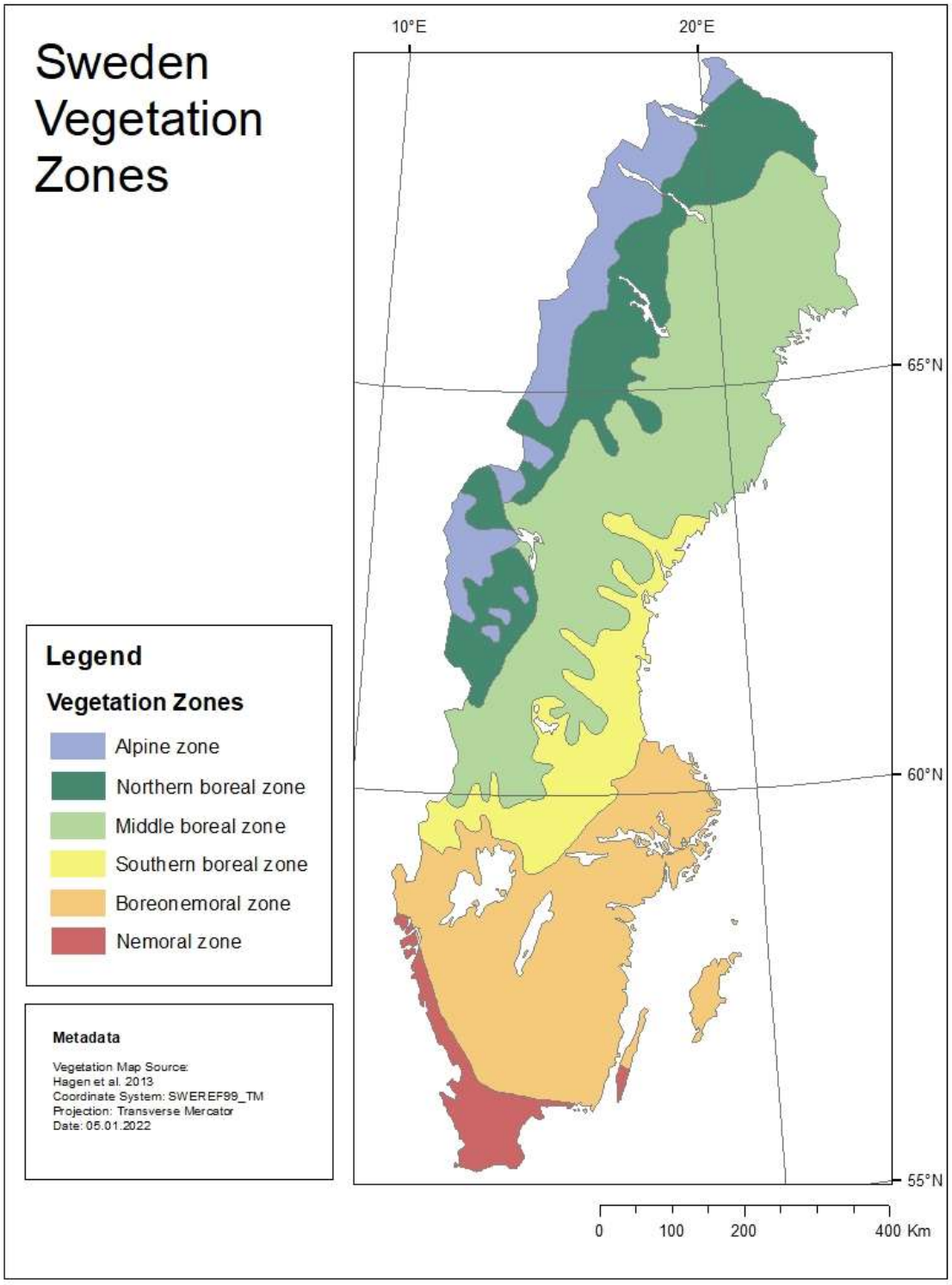


***Figure 1**: Vegetation zones in Sweden* (Hagen et al. 2013).

2.3. GIS Data Processing

The table below is a summary of the data, their sources, resolution and the file types.

The fire datasets were each individually unsuitable for use in this analysis due to various limitations. To address this, information from MSB fire statistics and EFFIS burn area maps were used to augment Skogsstyrelsen Burn Area Maps.

***Table 1**: The data and sources to be used in the analyses.*

Data	File type	Sources	Scale/Resolution
Fire statistics	.csv	MSB	Points (coordinates)
Skogsstyrelsen Burn area Map	Shapefile	Skogsstyrelsen	Fires >0.5ha mapped
EFFIS Burn area Map	Shapefile	EFFIS- Copernicus	Accurate >30ha
SLU Forest Maps	Raster-GeoTIFF	SLU	25m * 25m
Clearcut Map	Shapefile	Skogsstyrelsen	-
Road Map	Shapefile	Trafikverket	-
Rail Map	Shapefile	Lantmäteriet GSD	-
Water Map	Shapefile	Lantmäteriet GSD	-

2.3.1. Fire- Data Processing and Map Generation

The Skogsstyrelsen burn area maps have the finest resolution of any fire product covering the whole of Sweden (Skogsstyrelsen 2020). The delineated burn areas were produced by applying the Burn Area Index BAI2 to Sentinel-2 images from before and after each fire. The data covers fires from the 2018 to 2020. The dates that fires occured are missing for most fires.

The EFFIS burn area map is of coarse spatial resolution and lower accuracy than Skogsstyrelsen's, but includes fire occurrence dates (2009-2020), spatial resolution and accuracy improves in later years (2019 and 2020) (European Forest Fire Information System 2020).

MSB's (Myndigheten för samhällsskydd och Beredskap 2020) fire statistics data consists of the location, date, land use and area of fire events in Sweden from 1998-2020. The original dataset includes all fire events .

Skogsstyrelsen and EFFIS burn area maps were inspected and processed, cleaned to delete errors or duplicate features and obvious errors in internal shapefile vertices. Errors were detected manually through comparison of the three datasets, fires that were only supported by one dataset were excluded.

The dates of the fire events were obtained by cross referencing across all three fire datasets to produce a new dataset consisting of burn area maps for each year, with both high spatial and temporal resolution. Fires below 10ha were excluded, to limit the inclusion of fires which were heavily supressed. The workflow can be seen in Figure 3, where the burn area maps are called *Burn Map 2018*, 2019 and 2020 respectively. This dataset will be made availble for use by MSB, Skogsstyrelsen and SLU.

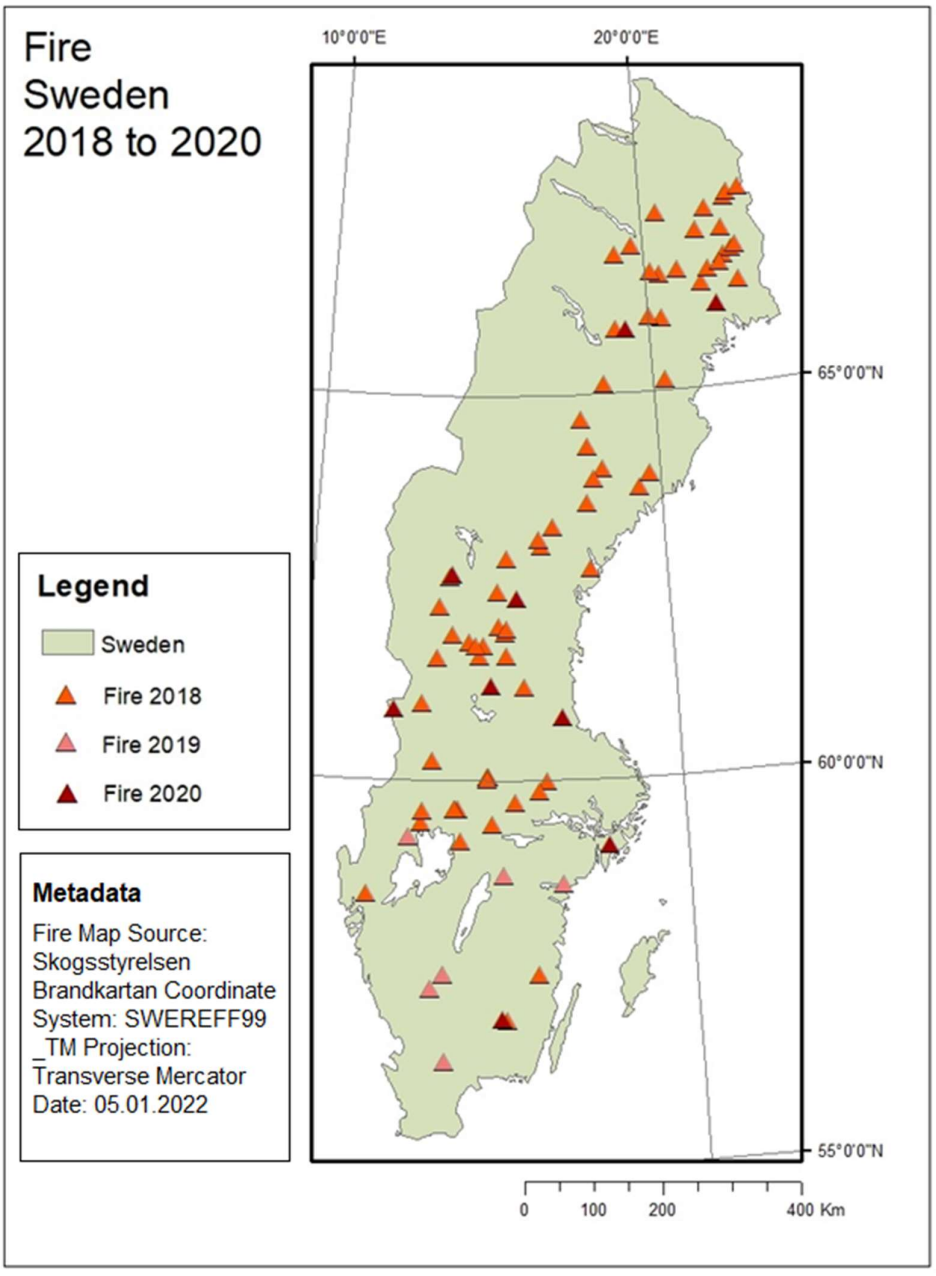


***Figure 2**: Map of fires >10ha in Sweden from 2018-2020.*

2.3.2. Firebreak- Data Processing and Map Generation

Roads, railways and waterbodies form pre-existing firebreaks. A firebreak map was created to exclude these firebreaks from sampling areas and limit their influence as firebreaks as much as possible. Firebreaks set up by firefighters are not recorded in any dataset, therefore could not be included.

Road shapefiles from Trafikverket (Trafikverket) and water and rail shapefiles from Lantmäteriet overview maps (Swedish National Land Survey 2021) were buffered by 10m either side, appended and exported as ***Firebreak map*** (see workflow diagram in Figure 3 below, and section "Preparing sampling areas" Figure 5).

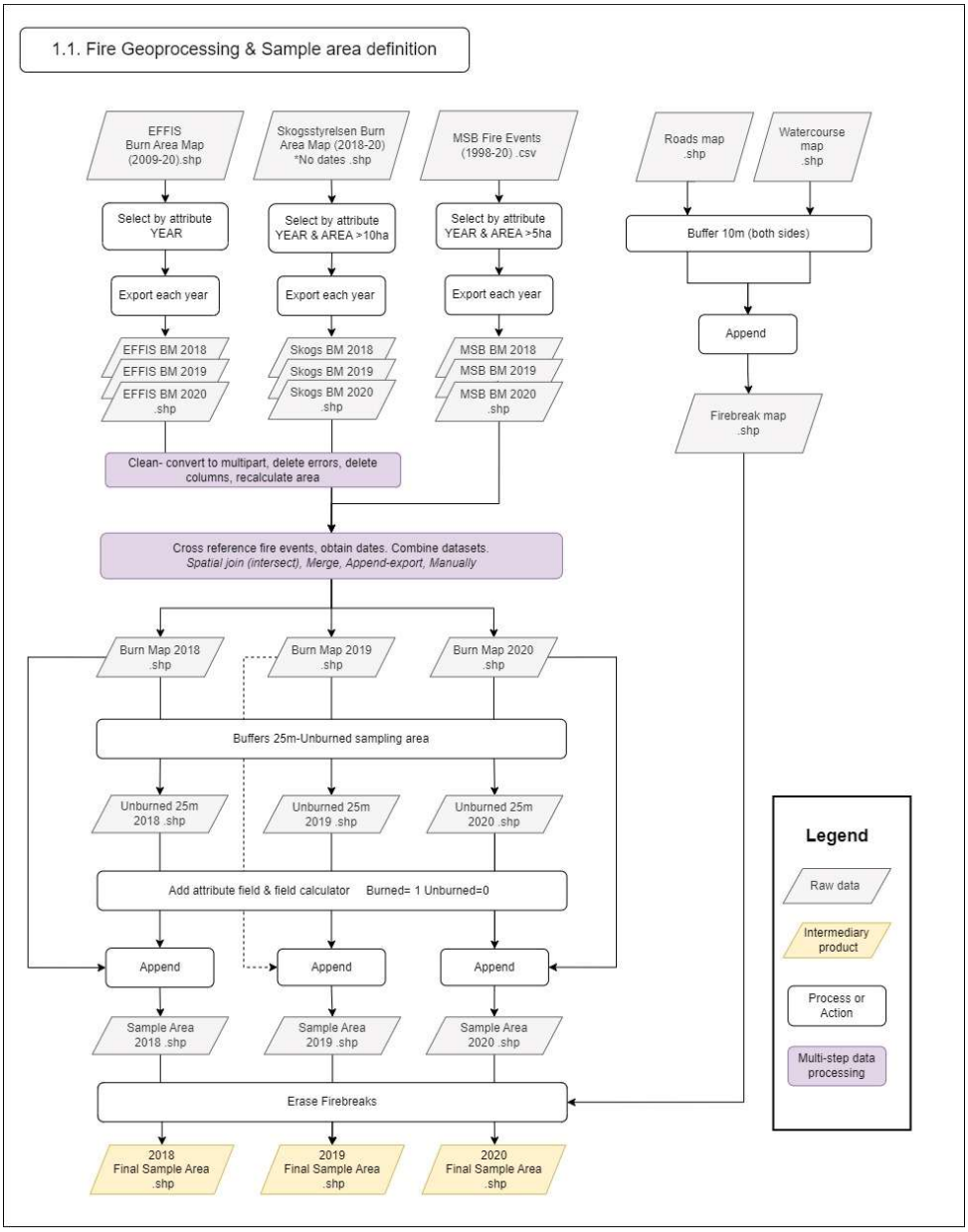


Figure 3*: Workflow showing the processing of GIS data and defining sample areas*

2.3.3. Forest -Data Processing and Map Generation

Maps of forest tree species present prior to fire events were taken from the SLU Forest Map 2010 (SLU Skogskartan, previous known as k-NN Sweden) (SLU 2010), from the Department of Forest Resource Management, Swedish University of Agricultural Sciences. SLU Forest Map 2010 used SPOT 4 and SPOT 5 imagery and National Forest Inventory field data, extrapolating estimates of forest parameter values using the k-nearest neighbour algorithm.

The Forest Map dataset used consisted of 10 individual GeoTIFFs covering the whole of Sweden, with a pixel size of 25x25m depicting the following variables: age of the stand, height, biomass (t ha^{-1}) and standing biomass volume per unit area (m^3sk ha^{-1}) for each of the following: Beech, Birch, Oak, Other deciduous, Pine, Spruce and Total volume (SLU 2010; Department of Forest Resource Management; SLU 2015). Standing biomass volume per hectare has been referred to as "volume" throughout the text for ease of reading.

Unfortunately, this method was discontinued after 2010, and subsequent years of SLU Skogskarta only map the Southern extent of Sweden, whereas most of the fires occur in the North.

***Table 2**: Forest Map variables and their units.*

Variable group	Variable name	Unit
Tree species	Scots Pine	m^3sk ha^{-1}
Tree species	Norway Spruce	m^3sk ha^{-1}
Tree species	Oak	m^3sk ha^{-1}
Tree species	Birch	m^3sk ha^{-1}
Tree species	Beech	m^3sk ha^{-1}
Vegetation parameters	Stand age	Years
Vegetation parameters	Height	m
Vegetation parameters	Biomass	t ha^{-1}
Vegetation parameters	Deciduous volume	m^3sk ha^{-1}
Vegetation parameters	Total volume	m^3sk ha^{-1}

Forest Maps were amended to account for clearcuts that occurred after 2010 and before the fires in 2018, 2019 and 2020, using the clearcut map from Skogsstyrelsen (Skogsstyrelsen 2021). The clearcut map (2003-2020) was derived from clearcuts registered with Skogsstyrelsen and before and after satellite images.

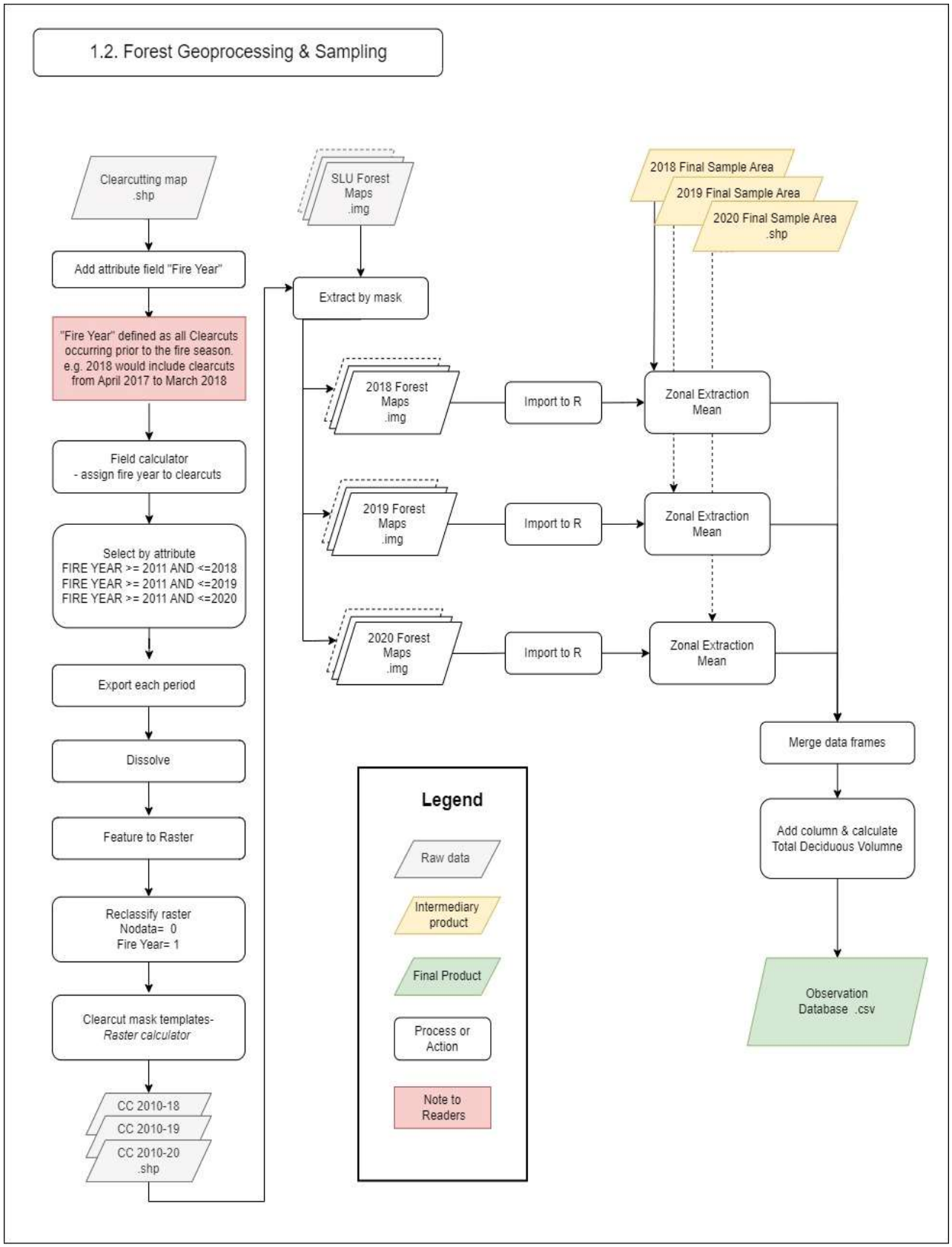


Figure 4: *Workflow outlining the geoprocessing of SLU forest map rasters and sampling.*

To do this, three forest map datasets were prepared, one for each fire year amended to account for clearcuts which occurred before the start of that year's fire season. The forest map dataset for 2018 was updated with all clearcuts which occurred between 2010 and March 31st 2018, the forest map dataset for 2019 included clearcuts from 2010-March 31st 2019 and forest map dataset 2020 those from 2010- March 31st 2020. First, clearcuts that occurred between 2010-March 31st 2018, 2010- March 31st 2019 and 2010-March 31st 2020 were exported to 3 new shapefiles and converted to raster. The DN values of clearcut cells were then reclassified to a value of 1.

Raster masks for each time period were created using raster calculator, each clearcut raster was added to a raster of Sweden (DN=1), and cells with a DN value of 2 were reclassified to "no data".

Each raster mask was applied to each of the 10 forest map rasters using "extract by mask". This resulted in a set of 10 forest map rasters for each fire year (2018, 2019, 2020) amended to exclude clearcuts which occurred before that year's fire season. The workflow for this data processing is outlined in Figure 4. These rasters were used for the extraction of explanatory variables.

2.4. Sampling and Data Extraction

Sample design posed several challenges, specifically regarding how to sample fire absence while accounting for extreme stochasticity, maintaining spatial independence of presence and absence locations and without altering the original research question. Sampling absence of fires from a buffer directly around the edge of the fires was the best approximation of absence of fire. Possible, advantages and disadvantages of this sampling method will be addressed in the discussion.

2.4.1. Sampling methods

The response variable was binary burned (1)/ not burned (0) and explanatory variables were taken as the mean values over the whole of each sample area. Each fire event represented one observation of fire presence, and the mean values of each explanatory variable for the entire area burned made up one row of the analysis database.

Observations of the absence of fire (not burned) were taken from a 25m buffer zone of unburned land around the perimeter of the fires. In this way each observation of presence was associated with an observation of absence, taken from the border where the fire stopped. The mean values of the explanatory variables for the entire area of an observed absence of fire also made up one row of the analysis database. This is illustrated in Figure 5 below.

2.4.2. Preparing Sampling Areas

To prepare the "unburned" areas from which to sample absence of fires the *Burn Maps* (2018, 2019 and 2020) were buffered by 25m (named e.g., *Unburned 25m 2018.shp*). An attribute was added to each buffer shapefile indicating the value of response variable (0). The shapefiles of unburned areas for each year were then appended to their associated *Burn Maps* and exported to *Sample Area* maps (one for each year) and the empty attribute field for response variable was filled as 1 for burned areas.

Figure 5*: Diagram illustrating the application of the firebreak map to sampling areas. The fire scar is shown in grey, and represents one observation of the burned class. The dark green buffer zone around the fire is the unburned sample area, representing one observation of the unburned class. Explanatory variables were extracted from each sample area; the mean value of all pixels within the sample area, for each forest raster.*

To limit the effects of firebreaks as much as possible (roads, railways and waterbodies) the *Firebreak map* was then erased from the *Sample Area* maps for each year, creating *Final Sample Area* maps for 2018, 2019 and 2020 (see workflow diagram in Figure 3 and illustration in Figure 5).

2.4.3. Sampling

The ***Final Sample Area*** maps and the ***Forest map*** rasters were imported to R. The mean values of the forest maps were extracted over the final sample areas. For example**,** ***Final Sample Area 2018*** contained polygons of burned areas (observations of presence) and polygons of unburned areas (observations of absence), for each of these polygons the mean of each forest variable was extracted from each of the 10 ***Forest Maps 2018***, see Figure 5 above. This was then stored as a data frame, where each row represents an observation of either presence or absence, with one column for each forest variable. This was then carried out for 2019 and 2020. The three data frames were then merged.

2.5 Analysis

Descriptive statistics was used to explore the data and an additional variable created by pooling all deciduous species volumes into total deciduous volume (m^3sk ha^{-1}). See Table 3 on the following page for a description of each explanatory variable.

***Table 3** : Explanatory variables reference guide*

Explanatory Variable Name	Expanded name	Unit	Description
Height	Mean tree height	m	Mean height of trees over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Age	Mean tree age	Years	Mean age of trees over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Biomass	Mean biomass above ground	$t\ ha^{-1}$	Mean biomass above ground over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Total Tree Volume	Mean total tree volume	$m^3sk\ ha^{-1}$	Mean total volume of all trees over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Pine	Mean pine volume	$m^3sk\ ha^{-1}$	Mean volume of all pine trees over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Spruce	Mean spruce volume	$m^3sk\ ha^{-1}$	Mean volume of all spruce trees over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Total Deciduous	Total of the mean deciduous tree volumes	$m^3sk\ ha^{-1}$	Total of the mean volume of all deciduous tree variables (oak, beech, birch, other deciduous), over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Birch	Mean birch volume	$m^3sk\ ha^{-1}$	Mean volume of all birch trees over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Other Deciduous	Mean volume of other deciduous trees	$m^3sk\ ha^{-1}$	Mean volume of all other deciduous trees over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Beech	Mean volume of beech	$m^3sk\ ha^{-1}$	Mean volume of all beech trees over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)
Oak	Mean volume of oak	$m^3sk\ ha^{-1}$	Mean volume of all oak trees over entire sample area according to SLU skogskarta 2010 (excluding clearcuts since 2010 and road verges)

2.5.1. Comparing Classes- Hypothesis test for two populations

Hypothesis testing was used to assess dependence between the burned and unburned classes and establish whether the two classes were significantly different enough to be considered to come from two different populations.

Shapiro-Wilk normality test was used to test the assumption of normally distributed differences between the two classes (burned and unburned) for each variable.

Due to non-normal distribution the non-parametric Wilcoxon matched-pairs signed rank test was used to test if the medians of the paired samples of the burned and unburned classes are significantly different enough to be considered two separate populations, using a significance level of alpha = 0.05.

2.5.2. Set-up for modelling

Spatial collinearity was not assessed as the sampling strategy outlined above results in paired observations, where the burned class was sampled from the fire and the unburned class form the buffer zone of the fire. When these paired observations had the same locations, spatial collinearity was assumed to be high. The individual fires themselves were typically widely spatially dispersed.

Explanatory variables were tested for pairwise collinearity and multicollinearity with variance inflation factors (Rodrigues and Riva 2014; Pourghasemi et al. 2020). Variables with strong linear correlations were excluded from model training in order of least important to most important, until the VIF value of each variable was within acceptable limits (under 10).

The R package caret: Classification and Regression Training (Max et al. 2022) was used to partition the dataset into training (80%) and test (20%) datasets. The training method was defined as 5-fold cross validation. Although the general rule of thumb is to use 10-fold cross validation, I used 5-fold cross validation due to a relatively small sample size which has been used in similar studies such as by Kalantar *et al.*, 2020. The caret package was also used to train the models (Max et al. 2022).

2.5.3. GLM- Logistic regression Models

Generalized Linear Models define the relationship between the response and explanatory variables (González et al. 2006; Rodrigues and Riva 2014; Pourghasemi et al. 2020).

GLM's generalize a typical linear regression using a link function, allowing a regression to be performed on non-normally distributed variables (Pourghasemi et al. 2020). The response variable is binary (0 or 1) and has a binomial-Bernoulli distribution, therefore the type of GLM used in this thesis is a logistic regression which uses LOGIT link function.
GLMs are statistical models traditionally used when modelling the probability of fire occurrence, as seen in literature from González *et al.*, 2006; Rodrigues and Riva, 2014; Pourghasemi *et al.*, 2020 and Jain *et al.*, 2020. Logistic regression was chosen due to its ease of use and interpretation (González et al. 2006), it's suitability to the binary nature of the response variable, and it has precedent in the literature. Non-parametric Wilcoxon matched-pairs signed rank tests (see Section 2.5.1. Comparing Classes above) were carried out to investigate whether burned and unburned classes were significantly different enough to be considered independent (See results section).

The multiple logistic regression equation predicts the transformed response variable, rather than the raw value as in a multiple linear regression (Hosmer et al. 2013; McDonald 2014), as such the equation is as follows:

$$\ln\left[\frac{P_x}{1-P_x}\right] = \beta_0 + \beta_1\ X_1 + \beta_2\ X_2 + \beta_3\ X_3 \ldots\ldots.$$

Where P_x is the probability of success (response variable is 1), β_0 is the intercept, $\beta_1\ X_1$ is the slope of the first explanatory variable, $\beta_2\ X_2$ is the slope of the second explanatory variable.

Alternatively the equation can be written as shown below (Hosmer et al. 2013; McDonald 2014) :

$$Px = \frac{e^{\beta_0+\beta_1\ X_1+\beta_2\ X_2+\beta_3\ X_3\ldots}}{1+e^{\beta_0+\beta_1\ X_1+\beta_2\ X_2+\beta_3\ X_3\ldots}} = \frac{1}{1+e^{\beta_0+\beta_1\ X_1+\beta_2\ X_2+\beta_3\ X_3\ldots}}$$

The logistic regression was carried out, with the training data and 5-fold cross validation. The P values of each variable were inspected for significance and model performance measured using the 5-fold cross validation accuracy, Cohen's kappa and Akaike Information Criterion (Max et al. 2022).

In logistic regression the response variable (y) has been transformed, which means interpretation of the model estimate coefficients is different from a linear regression (Hosmer et al. 2013; McDonald 2014). In a linear regression the value of the estimate for explanatory variable X is the change in the response variable (Y) with a one unit increase in X, (all other explanatory variables are held constant), and this represents the slope (Hosmer et al. 2013; McDonald 2014). E.g., If explanatory variable was mean annual rainfall (m) and the response variable was fire size (ha), if rainfall has an estimate of -1.3 it would mean that for every increase in 1m mean annual rainfall would result in a decrease in fire size of 1.3ha.

In logistic regression the unit of the response variable is the log of the odds ratio, and therefore, the coefficient estimate of each variable (slope) represents the change in the log odds of the response variable with one unit change in that variable (Hosmer et al. 2013). Log of odds ratio ranges from + infinity to -infinity (Hosmer et al. 2013).
The estimates of the coefficients were converted to odds ratios (OR) by taking the exponential of the regression coefficient, which is in log of the odds ratios.

$$OR = e^{\beta_1}$$

Each model builds out the equation

$$\ln\left[\frac{P_x}{1-P_x}\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \ldots \ldots$$

or for Probability:

$$Px = \frac{e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3 \ldots}}{1+e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3 \ldots}} = \frac{1}{1+e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3 \ldots}}$$

2.5.4. GLM- Predictions, Validations & Accuracy

The accuracy of the models was then tested with the partitioned test data (20%). The models were applied to the test dataset and used to predict the response variable based on the values of the explanatory variables of each observation. Predicted values of burned/unburned were then compared with true values and assessed using accuracy scores, Cohen's Kappa, confusion matrices and AUC ROC curves.

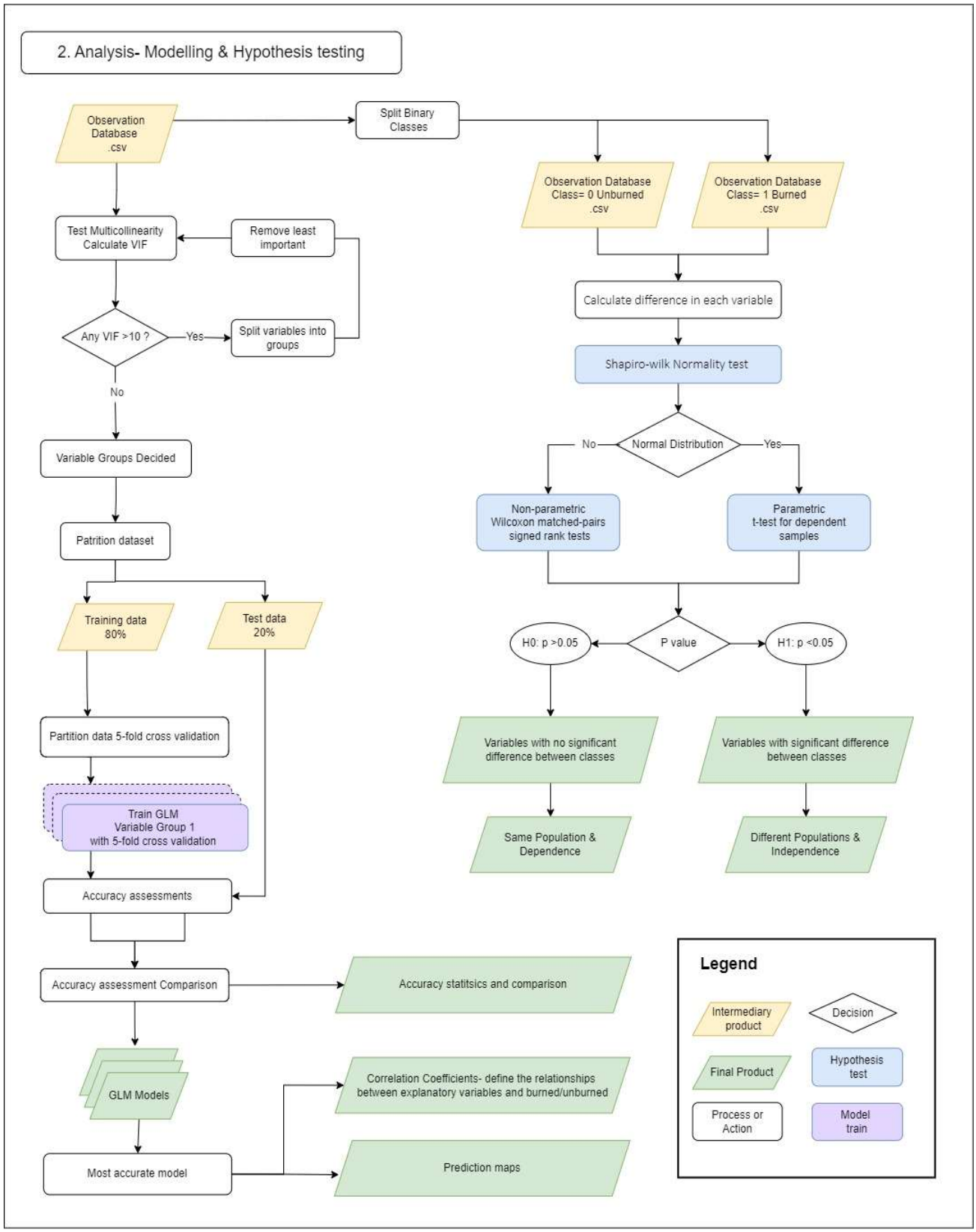


Figure 6: *Workflow outlining the analysis including modelling and hypothesis testing*

3. Results

3.1 Data Exploration

There was a total of 152 observations in the final dataset, 76 of which were burned observations and 76 were sampled unburned observations from the 25m buffer zone of the fire. Table 4 shows the mean value of explanatory variables over the entire burned area or fire-edge perimeter area.

Table 4: *Descriptive statistics of variables comparing burned and unburned observations*

Explanatory Variable	Total		Mean		Min		Max	
	Burned	Unburned	Burned	Unburned	Burned	Unburned	Burned	Unburned
Height (m)	850.872	812.572	11.196	10.692	2.804	2.875	18.091	17.182
Age (yrs)	5555.472	5141.417	73.098	67.650	18.795	22.860	161.046	158.147
Biomass ($t\ ha^{-1}$)	5910.802	5706.109	77.774	75.080	17.814	19.338	134.066	134.955
Total Tree Volume ($m^3sk\ ha^{-1}$)	8743.790	8398.506	115.050	110.507	21.873	23.150	215.343	216.364
Pine ($m^3sk\ ha^{-1}$)	5069.954	4374.403	66.710	57.558	10.382	10.000	158.794	117.796
Spruce ($m^3sk\ ha^{-1}$)	2697.289	2950.002	35.491	38.816	1.983	2.017	117.439	118.265
Total Deciduous ($m^3sk\ ha^{-1}$)	897.703	987.205	11.812	12.990	1.287	1.777	39.817	46.424
Birch ($m^3sk\ ha^{-1}$)	784.678	861.752	10.325	11.339	1.182	1.587	25.082	31.297
Other Deciduous ($m^3sk\ ha^{-1}$)	91.241	90.019	1.201	1.184	0.000	0.000	8.953	8.835
Beech ($m^3sk\ ha^{-1}$)	7.088	10.822	0.093	0.142	0.000	0.000	6.948	10.506
Oak ($m^3sk\ ha^{-1}$)	14.696	24.612	0.193	0.324	0.000	0.000	4.700	7.812

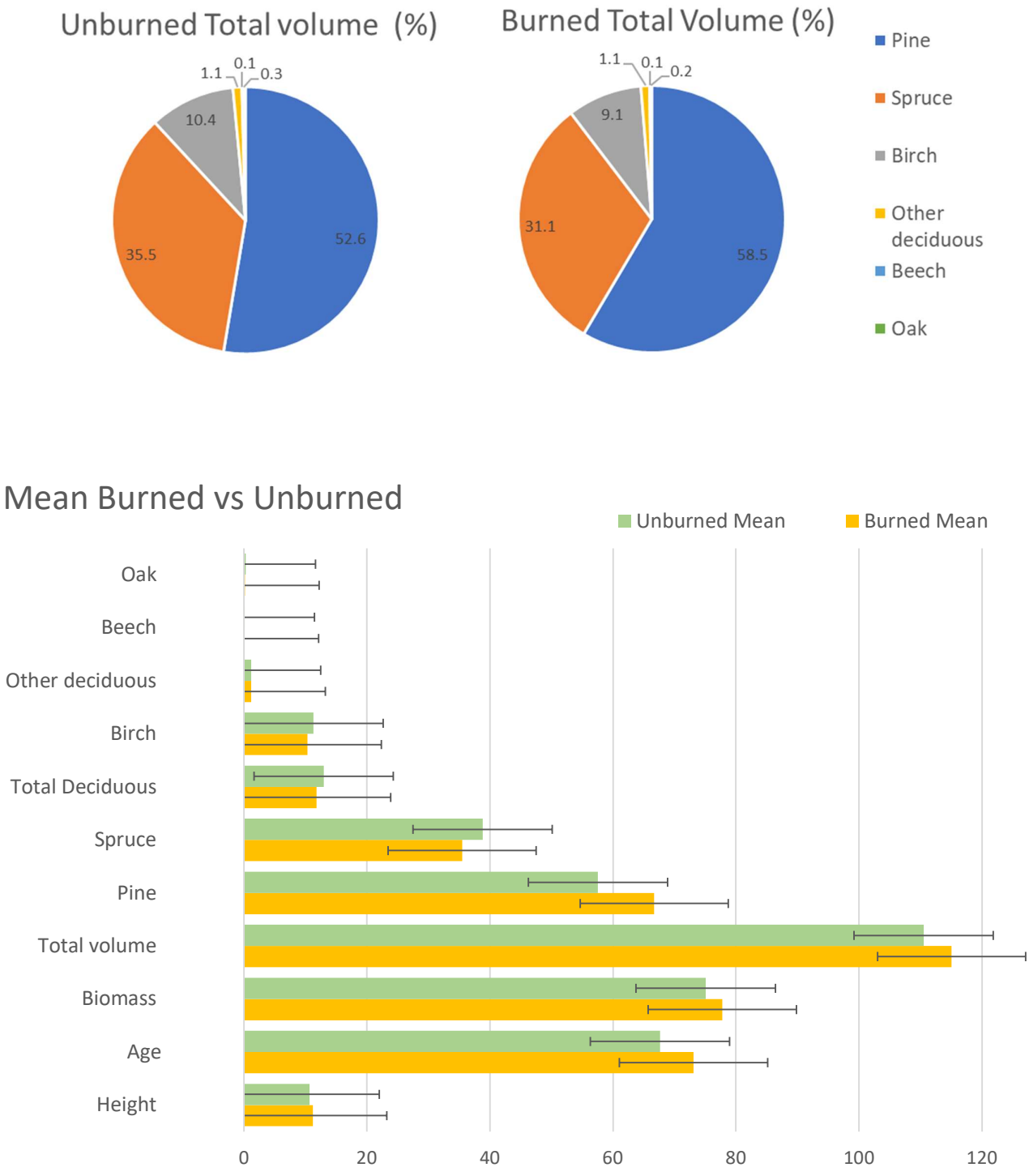


Figure 7: *Percentage total volume burned (top left) and unburned (top right) for each tree species and the variable distribution comparing means of explanatory variables for the burned and unburned classes (bottom).*

Tree height (m), age (years), biomass (t ha^{-1}), total tree volume ($m^3sk\ ha^{-1}$), pine ($m^3sk\ ha^{-1}$) and to a lesser extent other deciduous ($m^3sk\ ha^{-1}$), were all higher in the burned class (both the mean and total) (See Table 4 and Figure 7).

Spruce ($m^3sk\ ha^{-1}$), total deciduous ($m^3sk\ ha^{-1}$), birch ($m^3sk\ ha^{-1}$), beech ($m^3sk\ ha^{-1}$) and oak ($m^3sk\ ha^{-1}$) were all higher in the unburned class (both the mean and the total) (See Table 4 and Figure 7).

Beech, oak and to a lesser extent other deciduous were all underrepresented in the dataset.

Comparing the allocation of total volume in burned and unburned classes in Figure7 further highlights the differences in species composition of the two classes. In both classes, the majority of the volume was made up by pine followed by spruce, birch, other deciduous, beech and oak. Pine made up 58.5% of the total volume in burned observations, and 52.6% in unburned. Spruce made up a higher proportion of the total in unburned observations, with 31.1% in burned and 35.5% in unburned. Birch made up 9.1% of the total volume in burned observations and 10.4% in unburned.

The other deciduous tree species (beech, oak and other deciduous) were underrepresented in both classes and there was not much variation in their proportions, other than in oak which is 0.1% higher in unburned than burned.

The differences in distribution of each variable between burned and unburned observations are shown in the boxplots in Figure 7. The boxplots show the median, bounded by the interquartile range. The whiskers mark the range outside of which values can be considered outliers (1.5 IQR above and below the 3rd and 1st quartile respectively). Underrepresentation of oak and beech is well illustrated. Other deciduous species were also underrepresented, but not to the same extent.

The volume per hectare of birch, total deciduous and spruce was higher in the unburned class, whereas pine was higher in the burned class. Differences in distribution between classes looked to be inconclusive for other deciduous species, total volume.

Oak, beech and other deciduous were underrepresented in the dataset and zero inflated. There were differences in the means of all variables between classes, but it was not established whether these were significant.

3.2. Collinearity and Variance Inflation Factors

The pairwise collinearity is shown in Table 5 below using Pearson correlation coefficients. Values above 0.5 indicated collinearity and are highlighted.

Biomass was highly correlated with total tree volume, pine, spruce and height, biomass was not correlated with age or any of the deciduous variables, probably due to the lower representation of deciduous variables in the dataset. Total tree volume was also strongly correlated with pine, spruce and height, but showed no correlation with any of the deciduous variables. Both pine and spruce showed pairwise correlation with height. Total deciduous was the summed volume of all deciduous variables, as such it showed correlation with birch, oak and other deciduous, but not beech due to low values of beech.

The variable "other deciduous" (volume of other deciduous species) was correlated with birch and oak. Oak was correlated with beech. Age of tree showed no correlation with any other variable.

Variance inflation factors measure the correlation of one variable with all other variables. Values should be below 10 for all variables used together to train a model. Table 6 below illustrates the iterative removal of the most correlated variables.

Table 5. *Pairwise collinearity table showing the correlation coefficients of pairwise explanatory variables, values above 0.5 highlighted*

	Biomass	Age	Total Tree Volume	Pine	Spruce	Height	Birch	Other Deciduous	Oak	Beech	Total Deciduous
Biomass	1.000										
Age	0.151	1.000									
Total Tree Volume	0.990	0.051	1.000								
Pine	0.750	0.175	0.776	1.000							
Spruce	0.856	-0.063	0.858	0.370	1.000						
Height	0.937	0.358	0.918	0.791	0.723	1.000					
Birch	0.375	0.045	0.324	-0.112	0.420	0.285	1.000				
Other Deciduous	0.354	-0.296	0.362	0.016	0.403	0.219	0.515	1.000			
Oak	0.234	-0.151	0.226	0.010	0.245	0.197	0.250	0.572	1.000		
Beech	0.106	-0.128	0.108	-0.141	0.206	0.075	0.177	0.480	0.668	1.000	
Total Deciduous	0.408	-0.079	0.372	-0.095	0.468	0.300	0.915	0.768	0.557	0.479	1.000

Pairwise correlation coefficients and variance inflation factors were used to identify groups of variables with low collinearity which could be used to train the models. The aim was to limit collinearity while still testing relationships between the variables most interesting for the study. The groups of variables tested can be seen in Table 9 summarizing model results.

Table 6: *Variance Inflation Factors over several iterations of variable combinations. VIF should be below 10.*

Explanatory Variable	VIF Value			
	All	Remove derived variables	Remove biomass	Remove Height
Biomass	207.541	108.383	-	-
Age	5.734	4.030	2.849	1.263
Total Tree Volume	777.482	-	-	-
Pine	151.785	39.757	8.352	1.497
Spruce	137.138	40.721	5.510	1.698
Height	9.626	19.302	19.302	-
Birch	Infinite	4.979	2.174	1.818
Other Deciduous	Infinite	2.497	2.328	2.322
Oak	Infinite	2.423	2.265	2.151
Beech	Infinite	2.046	2.043	2.017
Total Deciduous	Infinite	-	-	-

3.3 Comparing Different Populations- Hypothesis test for two populations

To establish whether differences between burned and unburned classes were significant, hypothesis testing was carried out.

To decide if either a parametric or non-parametric hypothesis test for two populations would be suitable, the differences between two classes for each variable needed to be tested for normal distribution.

The Shapiro-Wilk normality test was used to test the assumptions of normally distributed differences between the two classes.

Null hypothesis: $H_0: p > 0.05$ *the distribution was not significantly different from normal distribution (assume normal distribution)*

Alternative hypothesis: : $H_A: p < 0.05$ *the distribution was significantly different from the normal distribution (assume non-normal distribution).*

Table 7: *Results of Shapiro Wilk Normality Test of the difference between classes.*

Variable	W value	P-value	Normal distribution H0: p >0.05	Non-Normal distribution HA: p <0.05
Height	0.938	0.001113	Reject	Accept
Age	0.885	0.000005	Reject	Accept
Biomass	0.915	0.000080	Reject	Accept
Total Tree Volume	0.923	0.000190	Reject	Accept
Pine	0.936	0.000851	Reject	Accept
Spruce	0.907	0.000036	Reject	Accept
Total Deciduous Volume	0.883	0.000004	Reject	Accept
Birch	0.953	0.006393	Reject	Accept
Other Deciduous	0.670	0.000000	Reject	Accept
Beech	0.103	< 2.2e-16	Reject	Accept
Oak	0.289	< 2.2e-16	Reject	Accept

None of the differences between the burned and unburned class, for any variable, had a p value above 0.05 indicating that all distributions were significantly different from the normal distribution. The null hypotheses were rejected as there was no evidence of normality.

Due to non-normal distribution the non-parametric Wilcoxon matched-pairs signed rank test was used to test if the medians of the paired samples of the burned and unburned classes are significantly different enough to be considered two separate populations, using a significance level of alpha = 0.05.

Null hypothesis: *H0: p > 0.05 there was no significant difference between the medians of the two classes.*

Alternative hypothesis: : *HA: p < 0.05 there was a significant difference between the medians of the two classes.*

Height (p=0.0370), age (p=0.0008), pine (p<1.2 *10^-5), total deciduous volume (p=0.0085) and birch (p=0.0089) had p-values under 0.05, indicating that there was a significant difference between the medians of the burned and unburned classes. However, once corrected for multiple testing using Holm-Bonferroni adjusted p-values height was no longer significant (adjusted p= 0.256900), total deciduous volume and birch volume were only marginally significant (adjusted p values of 0.076239 for both); therefore, the null hypothesis was accepted, there was no significant difference between the medians of the burned and unburned classes for these variables.

Age had an adjusted p value of 0.008013 and pine 0.000132, highlighting that there was a significant difference between the burned and unburned classes, the null hypothesis was rejected and the alternative hypothesis was accepted (see Table 8 below).

Biomass, total tree volume, spruce, other deciduous, oak and birch all had p-values above the alpha significance level 0.05, therefore the null hypothesis was accepted that there was no significant difference between the medians of the burned and unburned classes for these variables (see Table 8 below).

***Table 8**: Results of Wilcoxon matched-pairs signed rank test, testing whether samples are from different populations, with Holm-Bonferroni adjusted p-values accounting for multiple tests*

Variable	P-value	Adjusted P-value (Holm-Bonferroni)	No significant difference- Same Population H_0: $p > 0.05$	Significant difference- Different populations H_A: $p < 0.05$
Height	0.036700	0.256900	Accept	Reject
Age	0.000801	0.008013	Reject	Accept
Biomass	0.161400	0.718800	Accept	Reject
Total Tree Volume	0.123500	0.718800	Accept	Reject
Pine	0.000012	0.000132	Reject	Accept
Spruce	0.119800	0.718800	Accept	Reject
Total Deciduous Volume	0.008471	0.076239	Accept	Reject
Birch	0.008867	0.076239	Accept	Reject
Other Deciduous	0.615500	0.919800	Accept	Reject
Oak	0.306600	0.919800	Accept	Reject
Beech	0.422700	0.919800	Accept	Reject

These significant differences between the paired classes are illustrated in the frequency distribution histograms of delta (burned – unburned) for each variable below, including those variables with only marginal significance after correction for multiple testing (Figures 8-11). Negative delta values are those pairs in which values were higher in the unburned class, and positive delta values are from those pairs in which values were higher in the burned class.

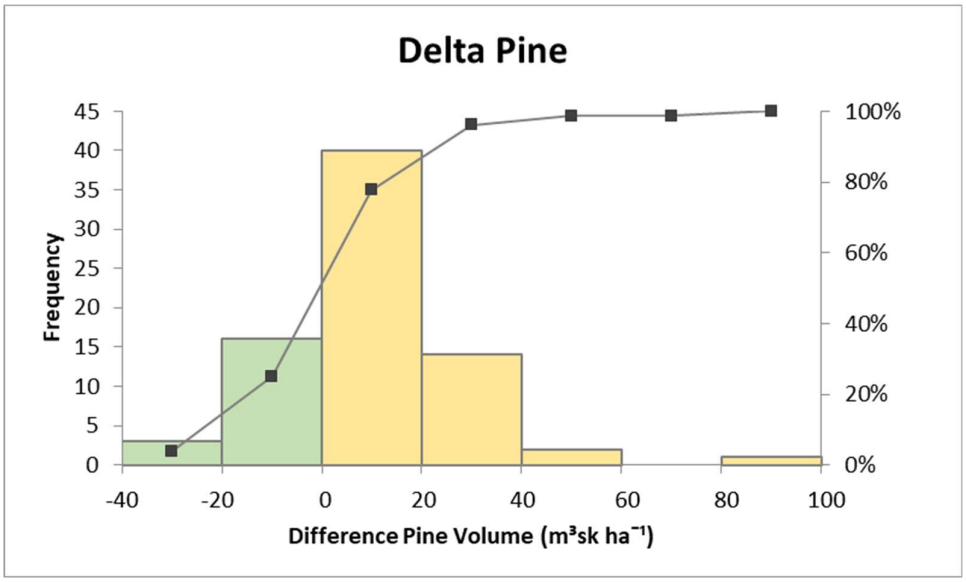


Figure 8: *Pine frequency distribution histogram showing the distribution of the differences between paired burned and unburned observations. (Burned-unburned), negative values are paired samples where the volume in the unburned observation was larger than the burned observation, positive values are from paired samples where to volume in burned was larger than unburned. 75% of pairs had higher pine volume in the burned observation.*

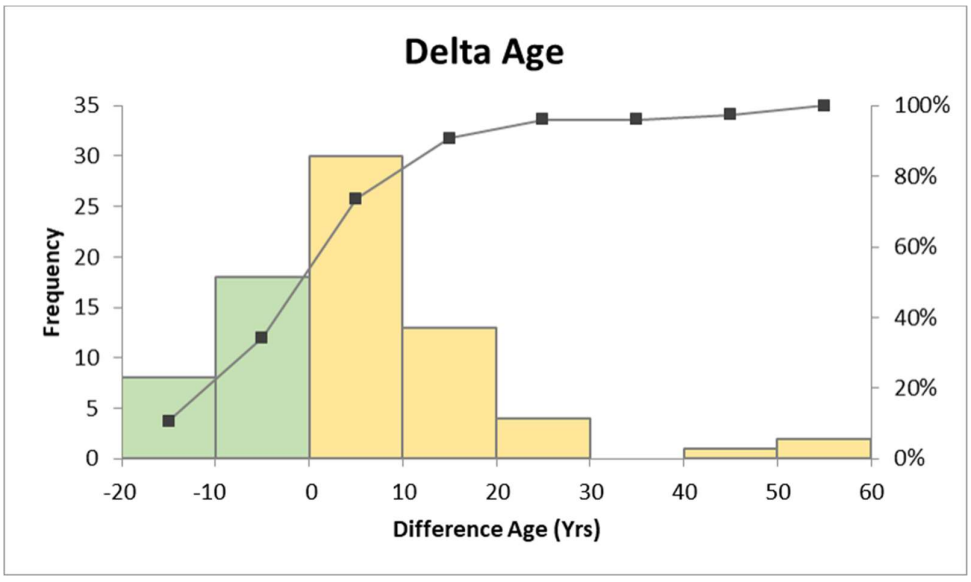


Figure 9: *Age frequency distribution histogram showing the distribution of the differences between paired burned and unburned observations. (Burned-unburned), negative values are paired samples where the volume in the unburned observation was larger than the burned observation, positive values are from paired samples where to volume in burned was larger than unburned. 65.79% of pairs had higher age in the burned observation.*

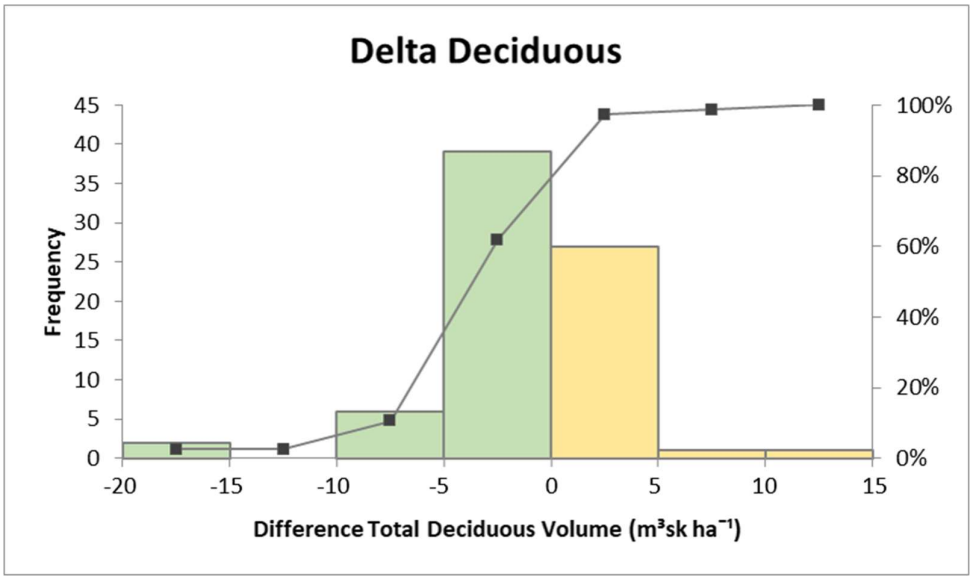


Figure 10*: Total deciduous frequency distribution histogram showing the distribution of the differences between paired burned and unburned observations. (Burned-unburned), negative values are paired samples where the volume in the unburned observation was larger than the burned observation, positive values are from paired samples where to volume in burned was larger than unburned. 61.84% of pairs had higher total deciduous volume in the unburned observation.*

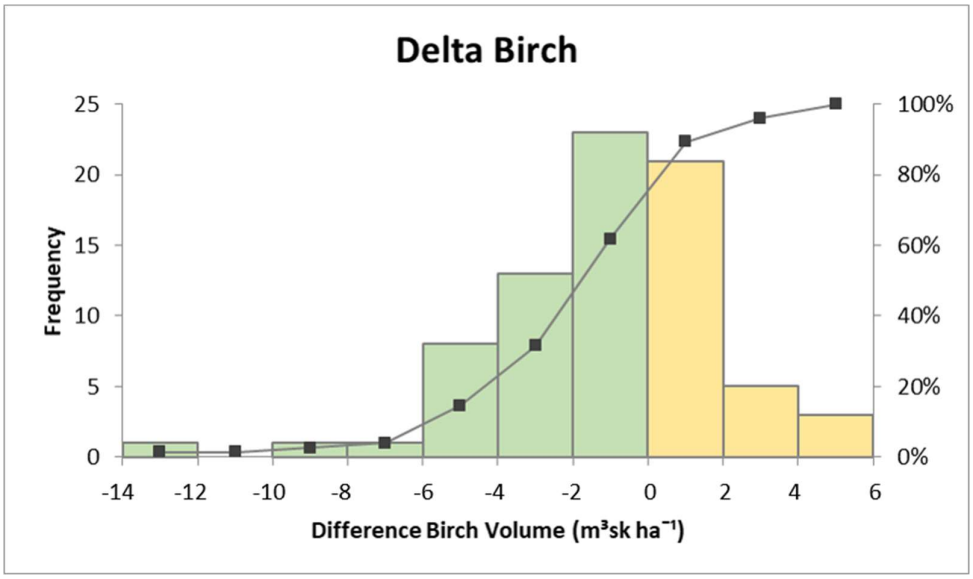


Figure 11: *Birch frequency distribution histogram showing the distribution of the differences between paired burned and unburned observations. (Burned-unburned), negative values are paired samples where the volume in the unburned observation was larger than the burned observation, positive values are from paired samples where to volume in burned was larger than unburned. 61.84% of pairs had higher total deciduous volume in the unburned observation.*

3.4. GLM Logistic Regression model results

Model Coefficients and Odds Plots

The logistic regression model estimates of coefficients and results are shown in Table 9 below.

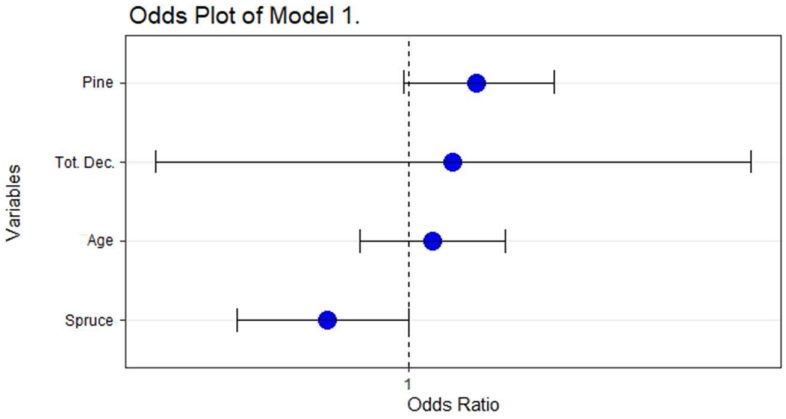
Table 9: *Coefficients of logistic regression models trained, including the estimate in probability.*

		Coefficients				
Model		Estimate	Std. Error	Z value	P value	Adjusted p-value (Holm-Bonferroni)
1.	Intercept	-0.698	0.761	-0.917	0.359	1.000
	Total Deciduous	0.009	0.031	0.293	0.769	1.000
	Pine	0.014	0.008	1.763	0.078	0.312
	Spruce	-0.017	0.009	-1.881	0.060	0.240
	Age	0.005	0.008	0.636	0.524	1.000
2.	Intercept	-0.286	0.474	-0.604	0.546	1.000
	Pine	0.014	0.008	1.878	0.060	0.121
	Spruce	-0.016	0.008	-2.124	0.034	0.067

Model 1.

The multiple logistic regression model 1 was trained assessing relationships between the response variable burned/unburned and the explanatory variables a.) mean total deciduous tree volume per hectare, b.) mean pine volume per hectare, c.) mean spruce volume per hectare and d). mean age.

The odds plot in Figure 12 below shows the estimates of the coefficients as odds ratios, indicated by the blue central point, the whisker bars on either side of each point indicate the upper and lower of the 95% confidence interval. An odds ratio of 1 represents a probability of 0.5 and is indicated by the dotted vertical line known as the "line of null effect." The smaller the confidence whiskers the lower the standard error. An odds ratio above 1 indicates a positive association between the explanatory variable and the response variable, and an odds ratio below 1 indicates a negative relationship. In this case a positive relationship would mean an increase in variable x results in an increase in relative probability of burning. If the confidence interval crosses the line of null effect then the standard error of the coefficient is too high for the variable to be considered significant.



Model 1.

Variable	Estimate	Confidence Interval	
	Odds Ratio	Low er	Upper
Tot. dec.	1.009	0.948	1.074
Pine	1.014	0.999	1.031
Spruce	0.983	0.965	1.000
Age	1.005	0.990	1.020

$$\ln\left[\frac{P_x}{1-P_x}\right] = -0.698 + 0.009(\text{tot. dec.}) + 0.014(\text{pine}) - 0.017(\text{spruce}) + 0.005(\text{age})$$

Figure 12: *The odds plot, odds ratios and equation of logistic regression Model 1.*

In Model 1, mean total deciduous volume had an estimate of 0.009 log of the odds ratio, thus an increase in 1 m^3sk ha^{-1} of total deciduous would result in an increase of 0.009 in the log of the odds ratio of the response variable being 1. Other than for identifying the direction of the relationship, interpretation of the raw log of odds ratios is not very intuitive, so instead I have used the odds ratios (as seen in Figure 12 above).

Total deciduous volume had an odds ratio estimate of 1.009, 95% CI(0.932, 1.091) indicating that holding all other variables constant, the odds of burning increased by 0.9% for each 1 m^3sk ha^{-1} increase of total deciduous volume (total deciduous volume ranged from ~1-46 m^3sk ha^{-1}). The p value for total deciduous was 0.769 (adjusted p value was close to 1), which is much greater than the threshold of 0.05, thus the relationship was not significant. This is clearly indicated in the odds plot (Figure 12), where the confidence interval extends widely and includes an odds ratio of 1.

Total deciduous p>0.05 the null hypothesis (b1=0)was accepted, there was no relationship between mean total deciduous volume and relative probability /likelihood of burning in Model 1.

Mean pine volume had an odds ratio estimate of 1.014, 95% CI(0.999, 1.031); holding all other variables constant, an increase in 1 m^3sk ha^{-1} of pine would result in an increased odds of burning of 1.4% (pine volume ranged from ~10-159 m^3sk ha^{-1}) . The confidence interval marginally extended past an odds ratio of 1, and the p value 0.078 was low, but not below the threshold of 0.05, and once corrected for multiple testing the Holm-Bonferroni adjusted p value was 0.312, was well above the threshold.

For pine p>0.05, and the null hypothesis (b1=0) was accepted. There was no relationship between mean pine volume and relative probability /likelihood of burning in Model 1.

Mean spruce volume had an odds ratio of 0.983, 95% CI(0.965, 1.000), an increase in 1

m^3sk ha^{-1} of spruce would result in a decrease in odds of burning of 1.7% (spruce volume ranged from ~2-118 m^3sk ha^{-1}). The confidence interval barely extended over the line of null effect and the p value of 0.06 reflected this almost significant relationship but it was not below the threshold of 0.05, once corrected for multiple testing the Holm-Bonferroni adjusted p value was 0.24, well above the threshold.

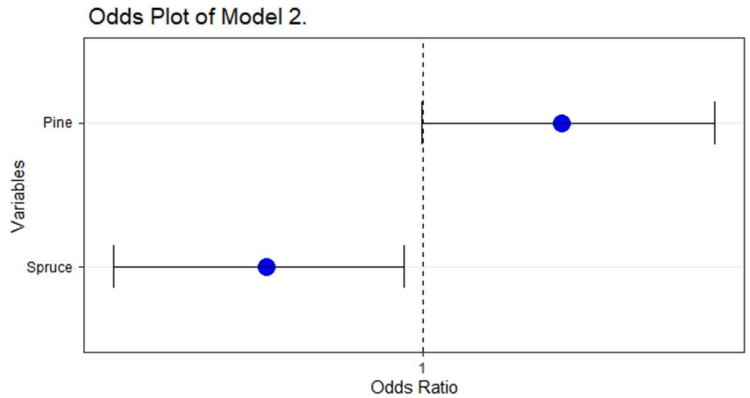
For spruce p>0.05, and the null hypothesis (b1=0) was accepted, as there was no detected relationship between mean spruce volume and relative probability /likelihood of burning in Model 1.

Mean age had a positive relationship with burning and a high p value (0.524, adjusted p value of 1), thus the null hypothesis was accepted, with age p>0.05 indicating there was no relationship between mean age and relative probability /likelihood of burning. As seen in the odds plot for Model 1, the odds ratio was 1.005, and confidence interval extends over the line of null effect.

In summary, pine had a positive relationship (1.4% increased risk), with a relatively low p value of 0.078, but high adjusted p value of 0.312 both above threshold of 0.05 to indicate significance. Spruce showed a negative relationship (1.7% decreased risk), with a relatively low p value of 0.06, but high adjusted p value of 0.24, neither were below the threshold of 0.05 needed for a significant relationship. Total deciduous volume had a positive relationship (0.9% increased risk) with a high p value of 0.769 (adjusted p value=1) with no significance. Age had a positive relationship (0.5% increased risk), with a high p value of 0.524 (adjusted p value=1) and therefore the relationship was not significant.

Model 2.

Multiple regression Model 2 was trained investigating the relationships between the response variable burned/unburned and the explanatory variables a.) mean pine volume per hectare and b.) mean spruce volume per hectare.



Model 2.

Variable	Estimate	Confidence Interval	
	Odds Ratio	Lower	Upper
Pine	1.014	1.000	1.030
Spruce	0.984	0.969	0.998

$$\ln\left[\frac{P_x}{1-P_x}\right] = -0.286 + 0.014(\text{pine}) - 0.016(\text{spruce})$$

Figure 13: *The odds plot, odds ratios and equation of logistic regression Model 2.*

Mean pine volume had an odds ratio estimate of 1.014, 95% CI (1.000, 1.030) making the odds of burning 1.441% more likely with each 1 m^3sk ha^{-1} increase in pine (pine volume ranged from ~10-159 m^3sk ha^{-1}). The p value of this relationship was 0.06, which was low, but not below the threshold of 0.05, and the adjusted p value was 0.121, indicating that this relationship was not significant. This is illustrated in the odds plot of Figure 13 above, where the confidence interval very marginally crosses the line of null effect.

For pine p>0.05, thus the null hypothesis (b1=0) was accepted, as there was no detectable relationship between mean pine volume and relative probability /likelihood of burning in Model 2.

Mean spruce volume had an odds ratio estimate of 0.984, 95% CI(0.969, 0.998), indicating that with each 1 m^3sk ha^{-1} increase in spruce the risk of burning is reduced by 1.585% (spruce volume ranged from ~2-118 m^3sk ha^{-1}). The confidence interval did not cross the odds ratio threshold of 1 (line of null effect) as seen in the odds plot above, illustrating the significance of the relationship, which had a p value of 0.034 which was below the threshold of 0.05. Once corrected for multiple testing the Holm-Bonferroni adjusted p value was 0.067, indicating marginal significance, but above the threshold of 0.05.

For spruce p>0.05, thus the null hypothesis (b1=0) was accepted, as there was no relationship between mean pine volume and relative probability /likelihood of burning in Model 2.

Model Accuracy

Model accuracy ranged from 0 to 1 with 0.5% indicating a 50% agreement between predicted and actual class values. Cohen's Kappa values define agreement as poor (under 0), slight (0-0.2), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80) and almost perfect (0.81-1.00) (Landis and Koch 1977).

The accuracy scores for both models were low to moderate during the cross validation, with Model 1 showing improved accuracy in final testing and Model 2 showing decreased accuracy in final testing.

Model 1 final testing resulted in an accuracy of 0.6, a kappa of 0.2 (slight to fair), and balanced sensitivity and specificity of 0.6.

In final testing, Model 2 showed a poor accuracy of 0.533, a kappa of 0.067 which could be considered as slight agreement between predicted and true observations. Sensitivity and specificity were unbalanced, where sensitivity was 0.6 and specificity 0.467 indicating fewer false negative and more false positives. See Table 10.

Table 10: *Accuracy statistics and significance of generalized linear models trained with different variable groups. Marginally significant P values highlighted in light green and significant in green. The blue cells highlight the best accuracy statistics.*

	Accuracy of 5-fold Cross Validation					Accuracy of final predictions			
	Accuracy	Kappa	AIC	Sig. Variables P value					
Model				Pine	Spruce	Accuracy	Kappa	Sens.	Spec.
Model 1									
Total Deciduous, Pine, Spruce, Age	0.516	0.032	172.100	0.078	0.060	0.600	0.200	0.600	0.600
Model 2									
Pine, Spruce	0.572	0.145	168.590	0.060	0.034	0.533	0.067	0.600	0.467

The receiver operating characteristic curves are shown in figure 14 below along with the area under curve (AUC) for each model. Probability thresholds were set to the standard 0.5, and not adjusted for either model as the number of observations of each response class (burned and unburned) were balanced. Model 1 had an AUC of 0.556 and Model 2 an AUC of 0.551, indicating a low ability to separate the burned class from the unburned class.

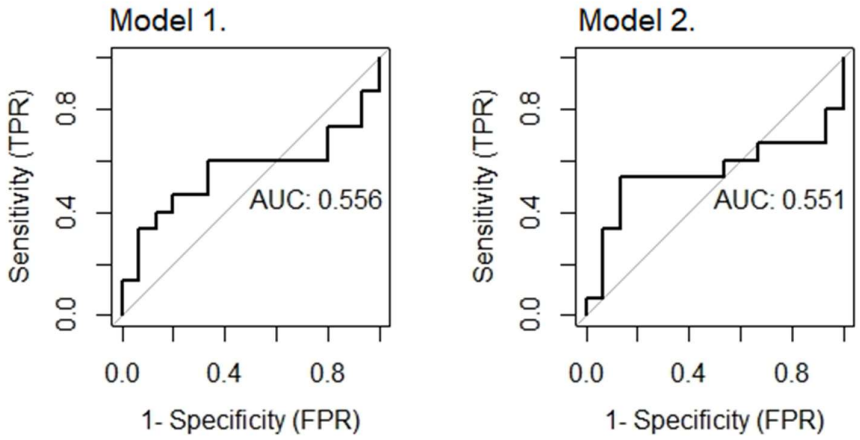


Figure 14*: AUC-ROC curves of Model 1 and Model 2.*

Predicted Probability Maps

Predicted maps of forest fire spread susceptibility using Model 1 are shown in Figures 15, 16 and 17.

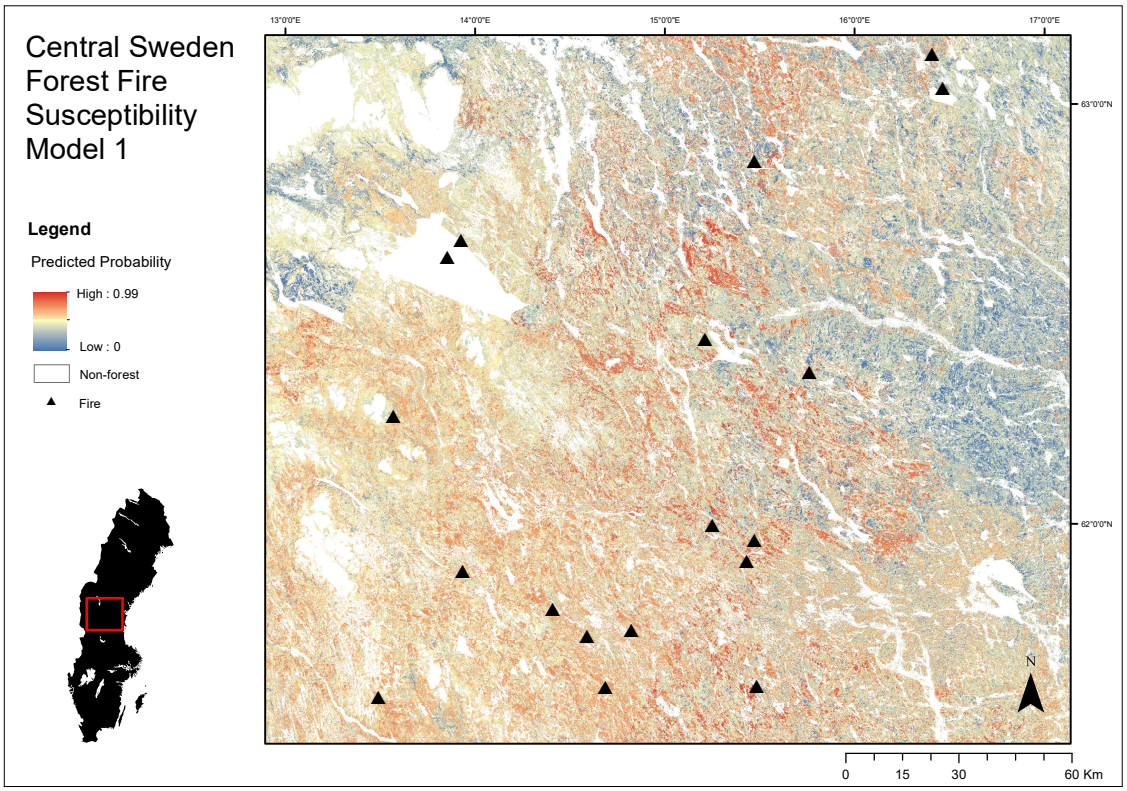


***Figure 15**: Map of predicted forest fire spread susceptibility (Model 1)of an area over central Sweden, and fire occurrence 2018-2020, including fires in non-forested areas.*

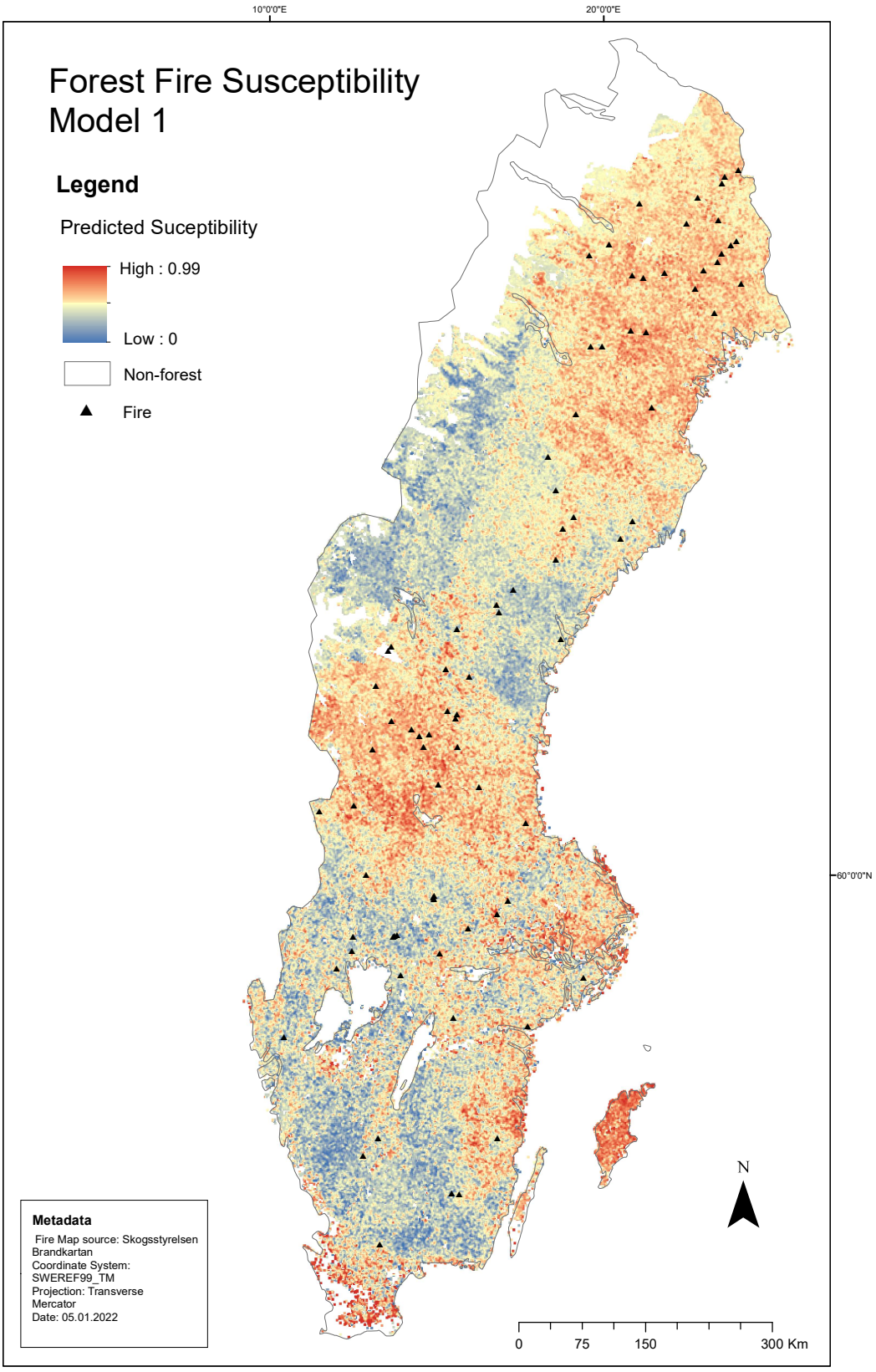


***Figure 16**: Mapped forest fire spread susceptibility of Sweden (Model 1)*

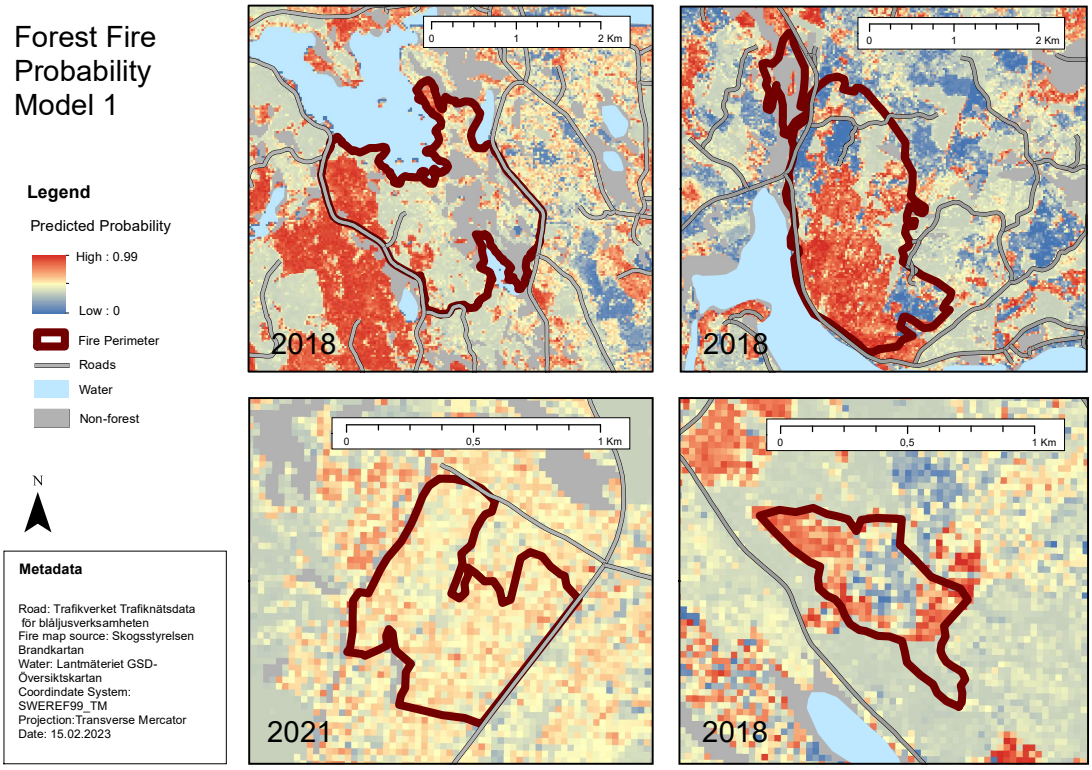


***Figure 17**: Map of predicted forest fire spread susceptibility of 4 selected areas in which fire occurred, 3 of which are fires which were part of the dataset (upper left, upper right and bottom right), and the map on the bottom left shows a fire from 2021 which was not used to train the model. Roads and water can clearly be seen to define many fire edges.*

4. Discussion

4.1 Limitations

Stochasticity & sampling
This thesis originally aimed to investigate the contribution of forest parameters to forest fire probability using large fires and existing forest parameter maps. However, fire occurrence is stochastic in nature, giving rise to some challenges, namely sampling fire absence while accounting for extreme stochasticity, while still addressing the original research question and maintaining spatial independence of presence and absence locations.

When modelling fire probability, there are three levels at which fire risk can be described: First is P1, the probability of ignition occurring and a small fire catching, second, P2, is the conditional probability of an ignition resulting in a large fire (spread), and the third, P3, is a product of P1 and P2, the unconditional probability of an ignition, fire starting, spreading and resulting in a large fire (Preisler et al. 2004; Preisler and Westerling 2007). In the large fires being investigated, both ignition and fire spread had already occurred, so to sample absence from a location with comparable conditions I sampled the fire edge. The fire edge is only location which received all the necessary preconditions needed to burn, and where the chance of burning has been maximized, this is the closest option for sampling true absence of fire as possible. In the end it was impossible to sample true absence and maintain spatial independence of presence and absence locations, and I prioritized true absence in the hopes of maximizing the signal of any patterns in the data.

Spatial Autocorrelation & paired samples
Models with high predictive power often have a greater reliance on spatial autocorrelation than the explanatory variables used to train the models (Ploton et al. 2020), but it is well known that fire propagates spatially. The most important factor in determining whether a location burns is whether it receives ignition. An existing fire provides ignition to all neighbouring areas, because fire is not static, but is both temporally and spatially dynamic. Fires move through both space and time. For this reason, I argue that, within the context of this study, trying to limit spatial autocorrelation would be an exercise in futility when modelling fire, as in the absence of spatial autocorrelation, fire would be static not spread.

Subsequently I modified the aims of the thesis to investigate forest fire spread susceptibility, rather than relative probability, and introduced the Wilcoxon matched-pairs signed rank test to test if the medians of the paired samples of the burned and unburned classes were significantly different enough to be considered two separate populations. This at least partially mitigates the worry that variables with significantly different classes violate the assumption of independent sampling for the generalized linear model.

SLU forest map 2010
The forest parameter data were from SLU forest maps from 2010, and the fires from 2018, 2019 and 2020. Unfortunately, SLU forest maps from 2015 only mapped the

Southern extent of Sweden where very few fires occurred. Additionally, the forest maps are not a true representation of reality, but rather an approximation of each parameter developed using inventory field data, satellite imagery and extrapolation using the k-nn algorithm. Forest parameters should be viewed as a proxy of the variable they represent. Finally, SLU forest maps are known lower accuracy estimates of deciduous tree volume in comparison to coniferous (Holmström et al. 2017) this was further compounded by true scarcity of deciduous trees in the landscape.

Burned Area Maps

The original plan was to use MSB fire statistics data and coordinates to point sample fire occurrence, however it quickly became clear that the coordinates and information in this dataset are extremely unreliable. The decision was made to source other fire maps. The EFFIS burn area map was low resolution and incomplete and the Skogsstyrelsen burn area map lacked dates and other information. Combining these datasets required substantial pre-processing, which was very time consuming.

4.2. Study Outcomes

The Wilcoxon matched-pairs signed rank test (subsequently referred to as the hypothesis testing), was undertaken to identify which variables had significantly different populations between the burned and unburned classes. The logistic regressions were carried out to model the relationships between multiple carefully selected variables and forest fire spread susceptibility.

Although height was initially found to be significantly different between burned and unburned classes, after correcting for multiple testing heights, this was not found to be significantly different between classes. This could be due to species specific height attributes, such as variations in ladder fuels associated with different species. Height was also highly correlated with other variables and therefore not included as a variable in logistic regression models. Birch volume and total deciduous volume were found to have only a marginal significant difference between the burned and unburned classes; only total deciduous volume was included in logistic regression modelling, due to pairwise collinearity between the two variables.

Oak, beech and other deciduous species were underrepresented in the dataset and the vast majority of observations had a value of 0, therefore it is was unsurprising that the hypothesis testing found no significant differences between the two classes, but a lack of significance cannot be ruled out due to their underrepresentation in the dataset.

Unexpectedly, the variables identified by hypothesis testing as having significantly different median values between the burned and unburned classes were not those which were found to contribute significantly to forest fire spread susceptibility in logistic regression models, with the possible exception of pine volume.

Although age was found to be significantly different in burned and unburned classes and total deciduous volume was found to be marginally significantly different after correcting for multiple testing. Unexpectedly, the logistic regression model showed that there were no significant relationships between either variable and forest fire spread susceptibility. Age was found to be higher in the burned class in 65.79% of paired observations, and had a positive contribution to forest fire spread susceptibility with no significance due to a large confidence interval. Total deciduous volume was higher in the unburned class in 61.84% of cases, but was found to have a positive contribution to forest fire spread susceptibility with no significance. When comparing the logistic regression models, Model 2 was trained without age and total deciduous volume had lower accuracy, kappa, and specificity than Model 1. This suggests that total deciduous volume or age are more important in predicting forest fire spread susceptibility than indicated by the Model 1, or they may be correlated with other environmental variables not included in this analysis which contribute to forest fire spread susceptibility. It is also interesting that Model 2, trained without age and total deciduous volume had lower specificity, which is an increase in Type I errors, an increase in false positives and over-predicting the burned case.

Conversely, spruce volume was the only variable which contributed with marginal significance (after correcting for multiple testing) to any logistic regression models, while hypothesis testing found no significant difference between the burned and unburned classes for spruce. Spruce volume was consistently associated with decreased forest fire spread susceptibility, despite the hypothesis that coniferous species would increase forest fire spread susceptibility. It is possible this is due to spruce prevalence in moist soils, both naturally and in managed forests.

Pine volume was the only variable in which there was agreement in both methods. After correcting for multiple testing, a significant difference was found between the burned and unburned classes in hypothesis testing. Pine volume contributed positively with marginal significance to forest fire spread susceptibility in logistic regression models, although this did not hold up after correcting for multiple testing as the adjusted p value was well above the alpha threshold for both models. One reason that pine volume impact on forest fire spread susceptibility was found to be only marginally significant, and then not significant after adjusting the p value, could be that the relationship is not linear. It could be that lower volume pine forest denotes a younger stand, in which trees are shorter with more ladder fuels and therefore easier to burn, in stands with very high volume of pine trees are likely to be older and taller, with a lack of ladder fuels, rendering them more resilient and less likely to burn.

The differences between burned and unburned classes highlighted by hypothesis testing showed relative proportional difference for each fire event individually. Whereas the logistic regression models did not take paired observations of the burned and unburned class into account. This indicates that on an individual fire level the proportional differences between the fire and the unburned surroundings were important, whereas when the dataset was pooled, and modelled between fire events, fewer relationships were identified. I expect that as each observation pair were neither spatially nor temporally distinct and experienced the same weather conditions it may be more useful to examine fires individually, as well as train models which include

temperature and a drought index as explanatory variables.

The low predictive power of both logistic regression models is likely to be due to multiple factors. Firstly, the dataset was small with only 152 observations, this was further complicated by 122 of these observations coming from one year, 2018, which was classed as a large fire year due to high temperatures, low rainfall and long fire season. It is thought that extreme drought homogenizes fuel conditions across stand types and differences in tree species flammability would become less prevalent under drought conditions (personal communication with I. Drobyshev, 2021).

As mentioned previously the forest parameters are based on extrapolated estimates, in order to identify relationships between forest fire spread susceptibility and tree volume by species, higher accuracy and a true representation of reality would be needed. Forest parameter values were also from 2010, therefore were 8 to 10 years out of date.

Firebreaks such as roads and waterbodies are known to play an important role in fire suppression (Araya et al. 2016; Pinto et al. 2020b). In this study a firebreak buffer was created and excluded from sample areas, this can clearly be seen in Figure 17 where roads and water are shown defining the fire edge. It would have been more useful to include them as variables in the modelling, but unfortunately this was outside the timeframe and scope of this MSc. Thesis.

Finally, other environmental variables more traditionally associated with modelling forest fire spread susceptibility, such as slope, aspect, elevation, temperature, precipitation (Sachdeva et al. 2018; Pourghasemi et al. 2020; Piao et al. 2022) fire weather index , drought code (Terrier et al. 2013; Drobyshev et al. 2021), soil moisture content (Pourghasemi et al. 2020; Drobyshev et al. 2021), wind speed and direction (Ryan and Opperman 2013; Jain et al. 2020; Pourghasemi et al. 2020) etc. would probably have resulted in models with higher predictive power, and it is likely that some of these other variables with well documented contribution to forest fire spread susceptibility were at play within the models.

4.3 Future studies

There are several ways in which the current models should be improved. The first step would be to incorporate newly delineated fires over 10ha for the years 2016, 2017, 2021 and 2022 to increase the size of the dataset. The soon to be released SLU forest map 2020 would provide more recent and accurate set of forest parameter maps for explanatory variables, especially in combination with the use of the Nationella Marktäckedata – a landcover class map. The inclusion of fire weather indices, such as drought code would provide useful insight into the strength of drought's relationship with forest fire spread susceptibility, especially in the context of deciduous versus coniferous stands.

It must be noted that relationships between forest fire spread susceptibility and a forest parameter may not be linear, and this cannot be accurately modelled by logistic

regression. It would be a useful exercise to carry out the same study using different models, for example machine learning models such as boosted regression trees and random forest modelling.

An alternative follow-up study would be to remove fires from the large fire year 2018, and adjust the sampling strategy using multiple point sampling for each fire and pseudoabsence point sampling. This would be followed by testing for spatial autocorrelation and using a mixed-effects model to account for spatial autocorrelation. This would remove fires which occurred under extreme weather conditions, and still substantially increase the dataset.

5. Conclusions

This study used hypothesis testing and logistic regression modelling to investigate the contribution of forest parameters, such as tree species, on forest fire spread susceptibility.

In conclusion, the hypothesis testing showed that tree age and pine volume were significantly different between the burned and unburned classes and the delta of paired observations showed that both were higher in burned than unburned classes. Birch and total deciduous volume were only marginally significantly different between the burned and unburned classes (after correcting p values for multiple testing), the delta of paired observations showed that that birch and deciduous volume were higher in the unburned class.

Logistic regression models showed the only variable with a marginal significant contribution to forest fire spread susceptibility was spruce volume, with an increase in spruce volume resulting in a decreased forest fire spread susceptibility. Pine volume initially had a marginally significant contribution to increased forest fire spread susceptibility, however after adjusting p values for multiple testing the contribution proved not significant.

Model 1, trained with age, deciduous volume, pine and spruce volume showed higher predictive power and balanced sensitivity and specificity than Model 2 which was trained with just pine and spruce volume. Model 2 also had unbalanced sensitivity and specificity, with lower specificity resulting in overpredicting the burned class.

The disagreement in results between the two methods (hypothesis testing and logistic regression models) could be due to multiple factors, among them most importantly a small sample size.

More research is needed with better data and sampling methods to fully investigate the contribution of tree species to forest fire spread susceptibility in Sweden.

References

Ahti, T., L. Hämet-Ahti, and J. Jalas. 1968. Vegetation zones and their sections in northwestern Europe. *Annales Botanici Fennici* 5: 169–211.

Araya, Y. H., T. K. Remmel, and A. H. Perera. 2016. What governs the presence of residual vegetation in boreal wildfires? *Journal of Geographical Systems* 18: 159–181. doi:10.1007/s10109-016-0227-9.

Chuvieco, E., I. Aguado, S. Jurdao, M. L. Pettinari, M. Yebra, J. Salas, S. Hantson, J. De La Riva, et al. 2014. Integrating geospatial information into fire risk assessment. *International Journal of Wildland Fire* 23: 606–619. doi:10.1071/WF12052.

Department of Forest Resource Management; SLU. 2015. SLU Skogskarta Metadata: 2000–2015. Swedish University of Agricultural Sciences.

Drobyshev, I., M. Niklasson, and H. W. Linderholm. 2012. Forest fire activity in Sweden: Climatic controls and geographical patterns in 20th century. *Agricultural and Forest Meteorology* 154–155. Elsevier B.V.: 174. doi:10.1016/j.agrformet.2011.11.002.

Drobyshev, I., N. Ryzhkova, J. Eden, M. Kitenberga, G. Pinto, H. Lindberg, F. Krikken, M. Yermokhin, et al. 2021. Trends and patterns in annually burned forest areas and fire weather across the European boreal zone in the 20th and early 21st centuries. *Agricultural and Forest Meteorology* 306. Elsevier B.V.: 108467. doi:10.1016/j.agrformet.2021.108467.

ESRI. 2017. ArcGIS Desktop: release 10.5.1.

European Forest Fire Information System. 2020. EFFIS Burnt Area Product. European Forest Fire Information System -Copernicus.

Flannigan, M., A. S. Cantin, W. J. De Groot, M. Wotton, A. Newbery, and L. M. Gowman. 2013. Global wildland fire season severity in the 21st century. *Forest Ecology and Management* 294: 54–61. doi:10.1016/j.foreco.2012.10.022.

González, J. R., M. Palahi, Trasobares, Antoni, and T. Pukkala. 2006. A fire probability model for forest stands in Catalonia (north-east Spain) José. *Annals of Forest Science* 63. doi:10.1051/forest.

Hagen, D., K. Svavarsdottir, C. Nilsson, A. K. Tolvanen, K. Raulund-Rasmussen, À. L. Aradòttir, A. M. Fosaa, and G. Halldorsson. 2013. Ecological and social dimensions of ecosystem restoration in the nordic countries. *Ecology and Society* 18. doi:10.5751/ES-05891-180434.

Hély, C., Y. Bergeron, and M. D. Flannigan. 2000. Effects of Stand Composition on Fire Hazard in Mixed-Wood Canadian Boreal Forest. *Journal of Vegetation Science* 11: 813–824.

Hély, C., M. Flannigan, Y. Bergeron, and D. McRae. 2001. Role of vegetation and weather on fire behavior in the Canadian mixedwood boreal forest using two fire behavior prediction systems. *Canadian Journal of Forest Research* 31: 430–441. doi:10.1139/x00-192.

Hély, C., C. M. J. Fortin, K. R. Anderson, and Y. Bergeron. 2010. Landscape composition influences local pattern of fire size in the eastern Canadian boreal forest: Role of weather and landscape mosaic on fire size distribution in mixedwood boreal forest using the Prescribed Fire Analysis System. *International Journal of Wildland Fire* 19: 1099–1109. doi:10.1071/WF09112.

Holmström, E., M. Karlsson, and U. Nilsson. 2017. Modeling birch seed supply and seedling establishment during forest regeneration. *Ecological Modelling* 352. Elsevier B.V.: 31–39. doi:10.1016/j.ecolmodel.2017.02.027.

Hörnberg, G., J. E. Wallin, T. Påsse, D. A. Wardle, and O. Zackrisson. 2004. Holocene land uplift and its influence on fire history and ecosystem development in boreal Sweden. *Journal of Vegetation Science* 15: 171–180. doi:10.1111/j.1654-1103.2004.tb02252.x.

Hosmer, D. W., S. Lemeshow, and R. X. Sturdivant. 2013. *Applied Logistic Regression. Journal of the American Statistical Association*. Hoboken, New Jersey: John Wiley & Sons. doi:10.2307/2290035.

Jain, P., S. C. P. Coogan, S. G. Subramanian, M. Crowley, S. Taylor, and M. D. Flannigan. 2020. A review of machine learning applications in wildfire science and management. *Environmental Reviews* 28: 478–505. doi:10.1139/er-2020-0019.

Johnson, E. A., A. M. Gill, R. Bradstock, A. Granstrom, and L. Trabaud. 2003. Towards a sounder fire ecology. *Frontiers in Ecology and the Envirswedenonment* 1: 271–276. doi:10.2307/3868015.

Kalantar, B., N. Ueda, M. O. Idrees, S. Janizadeh, K. Ahmadi, and F. Shabani. 2020. Forest fire susceptibility prediction based on machine learning models with resampling algorithms on remote sensing data. *Remote Sensing* 12: 1–24. doi:10.3390/rs12223682.

Landis, J. R., and G. G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33: 159. doi:10.2307/2529310.

Max, A., J. Wing, S. Weston, A. Williams, C. Keefer, A. Engelhardt, T. Cooper, Z. Mayer, et al. 2022. caret: Classification and Regression Training. R package.

McDonald, J. H. 2014. Multiple logistic regression. In *Handbook of Biological Statistics*, 3rd ed., 231–239. Baltimore: Sparky House Publishing.

Molinari, C., C. Carcaillet, R. H. W. Bradshaw, G. E. Hannon, and V. Lehsten. 2020. Fire-vegetation interactions during the last 11,000 years in boreal and cold temperate forests of Fennoscandia. *Quaternary Science Reviews* 241. Elsevier Ltd: 106408. doi:10.1016/j.quascirev.2020.106408.

Myndigheten för samhällsskydd och Beredskap. 2020. Incident reports from municipal fire brigades. Myndigheten för samhällsskydd och Beredskap.

Niklasson, M., and A. Granström. 2000. Numbers and Sizes of Fires : Long-Term Spatially Explicit Fire History in a Swedish Boreal Landscape. *Ecology* 81: 1484–1499.

Nilsson, P., C. Roberge, J. Fridman, and S. Wulff. 2021. *Skogsdata 2021: Aktuella uppgifter om de svenska skogarna från SLU Riksskogstaxeringen*.

Päätalo, M.-L. 1998. Factors influencing occurrence and impacts of fires in northern European forests. *Silva Fennica* 32: 185–202.

Piao, Y., D. Lee, S. Park, H. G. Kim, and Y. Jin. 2022. Forest fire susceptibility assessment using google earth engine in Gangwon-do, Republic of Korea. *Geomatics, Natural Hazards and Risk* 13. Taylor & Francis: 432–450. doi:10.1080/19475705.2022.2030808.

Pinto, G., M. Niklasson, N. Ryzhkova, and I. Drobyshev. 2020a. A 500-year history of forest fires in Sala area, central Sweden, shows the earliest known onset of fire suppression in Scandinavia. *Regional Environmental Change* 20. Regional Environmental Change. doi:10.1007/s10113-020-01718-2.

Pinto, G., F. Rousseu, M. Niklasson, and I. Drobyshev. 2020b. Effects of human-related and biotic landscape features on the occurrence and size of modern forest fires in Sweden. *Agricultural and Forest Meteorology* 291. Elsevier: 108084. doi:10.1016/j.agrformet.2020.108084.

Ploton, P., F. Mortier, M. Réjou-Méchain, N. Barbier, N. Picard, V. Rossi, C. Dormann, G. Cornu, et al. 2020. Spatial validation reveals poor predictive performance of large-

scale ecological mapping models. *Nature Communications* 11. Springer US: 1–12. doi:10.1038/s41467-020-18321-y.
Pourghasemi, H. R., A. Gayen, R. Lasaponara, and J. P. Tiefenbacher. 2020. Application of learning vector quantization and different machine learning techniques to assessing forest fire influence factors and spatial modelling. *Environmental Research* 184. Elsevier Inc.: 109321. doi:10.1016/j.envres.2020.109321.
Preisler, H. K., and A. L. Westerling. 2007. Statistical model for forecasting monthly large wildfire events in Western United States. *Journal of Applied Meteorology and Climatology* 46: 1020–1030. doi:10.1175/JAM2513.1.
Preisler, H. K., D. R. Brillinger, R. E. Burgan, and J. W. Benoit. 2004. Probability based models for estimation of wildfire risk. *International Journal of Wildland Fire* 13: 133–142. doi:10.1071/WF02061.
Roberge, J.-M., C. Fries, E. Norwark, E. Mårald, A. Sténs, C. Sandström, J. Sonesson, C. Appelqvist, et al. 2020. *Forest management in Sweden Current practice and historical background*. Jönköping.
Rodrigues, M., and J. De Riva. 2014. Environmental Modelling & Software An insight into machine-learning algorithms to model human-caused wild fire occurrence. *Environmental Modelling and Software* 57. Elsevier Ltd: 192–201. doi:10.1016/j.envsoft.2014.03.003.
Ryan, K. C., and T. S. Opperman. 2013. LANDFIRE - A national vegetation/fuels data base for use in fuels treatment, restoration, and suppression planning. *Forest Ecology and Management* 294: 208–216. doi:10.1016/j.foreco.2012.11.003.
Sachdeva, S., T. Bhatia, and A. K. Verma. 2018. GIS-based evolutionary optimized Gradient Boosted Decision Trees for forest fire susceptibility mapping. *Natural Hazards* 92. Springer Netherlands: 1399–1418. doi:10.1007/s11069-018-3256-5.
Scott, J. H., and R. E. Burgan. 2005. *Standard Fire Behavior Fuel Models: A Comprehensive Set for Use with Rothermel's Surface Fire Spread Model*. *General Technical Report RMRS-GTR-153*. Fort Collins.
Sjöström, J., and A. Granström. 2020. *Skogsbränder och gräsbränder i Sverige - Trender och mönster under senare decennier*.
Skogsstyrelsen. 2020. Skogsbränder.
Skogsstyrelsen. 2021. Utförda avverkningar.
SLU. 2010. kNN-Sverige – Aktuella kartdata över skogsmarken, årgång 2005 och 2010 (in Swedish). Sveriges lantbruksuniversitet.
Swedish National Land Survey, (Lantmäteriet). 2021. GSD-General Map, vector format. Lantmäteriet.
Terrier, A., M. P. Girardin, C. Périé, P. Legendre, and Y. Bergeron. 2013. Potential changes in forest composition could reduce impacts of climate change on boreal wildfires. *Ecological Applications* 23: 21–35. doi:10.1890/12-0425.1.
The R Foundation for Statistical computing. 2020. R version 3.6.3.
Trafikverket. Trafiknätsdata för blåljusverksamheten. *[Map Data, Region Nord, Mitt, Väst, Öst, Stockholm, Syd]*.
Van Wagtendonk, J. W., J. M. Benedict, and W. M. Sydoriak. 1996. Physical properties of woody fuel particles of Sierra Nevada conifers. *International Journal of Wildland Fire* 6: 117–123. doi:10.1071/WF9960117.
Wastenson, L., V. Haldo, and B. Raab. 1995. *National atlas of Sweden: Climate, lakes and rivers*. First. Stockholm: Sveriges nationalatlas.
Xanthopoulos, G., C. Calfapietra, and P. Fernandes. 2012. Fire Hazard and Flammability of European Forest Types. In *Post-fire forest management in southern Europe: a*

COST action for gathering and disseminating scientific knowledge, ed. V. R. Vallejo, M. Arianoutsou, and F. Moreira, 79–92. doi:10.1007/978-94-007-2208-8.

Appendix

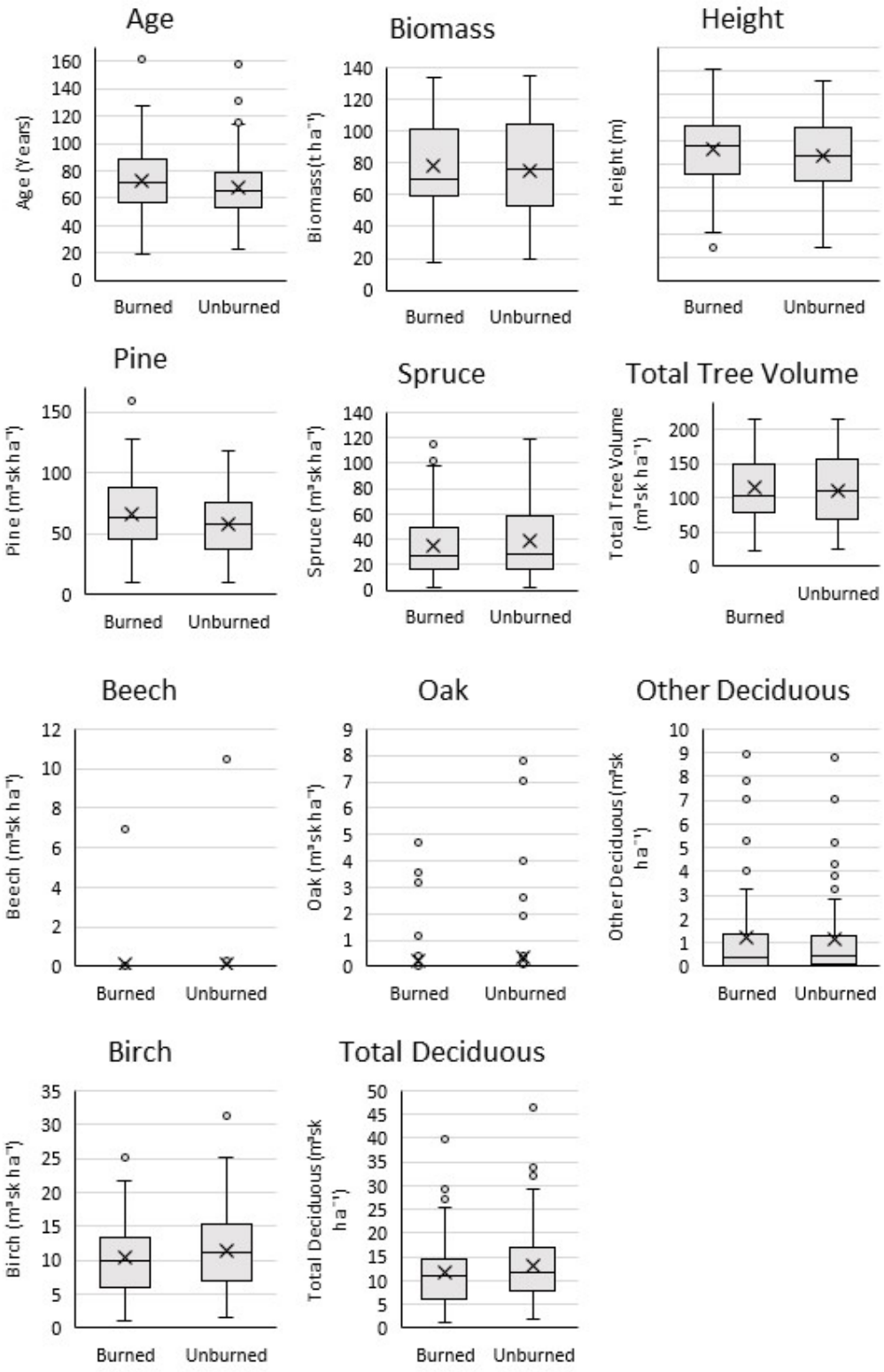


Figure 18: Variable distribution comparing the Burned and Unburned Classes

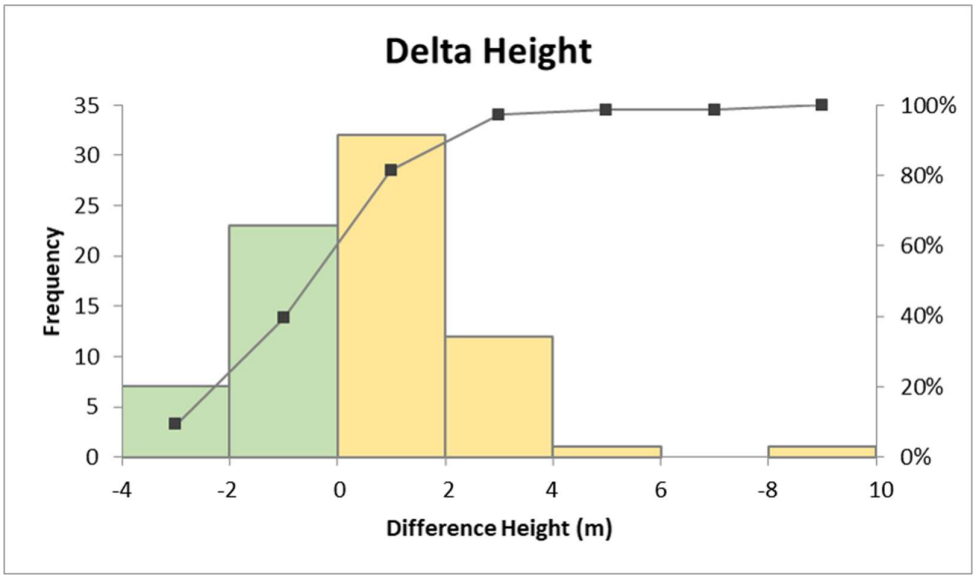


Figure 19*: Height frequency distribution histogram showing the distribution of the differences between paired burned and unburned observations. (Burned-unburned), negative values are paired samples where the volume in the unburned observation was larger than the burned observation, positive values are from paired samples where to volume in burned was larger than unburned. 60.53% of pairs had higher height in the burned observation.*

Series from Lund University

Department of Physical Geography and Ecosystem Science

Master Thesis in Geographical Information Science

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland (2008).
2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research (2008).
3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing (2009).
4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia (2009).
5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).
6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).
7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010).
8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010).
9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010).
10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation and malaria transmission intensities in selected parts of Africa (2011).
11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011).
12. *Jonathan Higgins:* Monitoring urban growth in greater Lagos: A case study using GIS to monitor the urban growth of Lagos 1990 - 2008 and produce future growth prospects for the city (2011).
13. *Mårten Karlberg:* Mobile Map Client API: Design and Implementation for Android (2011).
14. *Jeanette McBride:* Mapping Chicago area urban tree canopy using color infrared imagery (2011).
15. *Andrew Farina:* Exploring the relationship between land surface temperature and vegetation abundance for urban heat island mitigation in Seville, Spain (2011).
16. *David Kanyari*: Nairobi City Journey Planner: An online and a Mobile Application (2011).
17. *Laura V. Drews:* Multi-criteria GIS analysis for siting of small wind power plants - A case study from Berlin (2012).
18. *Qaisar Nadeem:* Best living neighborhood in the city - A GIS based multi criteria evaluation of ArRiyadh City (2012).
19. *Ahmed Mohamed El Saeid Mustafa:* Development of a photo voltaic building rooftop integration analysis tool for GIS for Dokki District, Cairo, Egypt (2012).

20. *Daniel Patrick Taylor*: Eastern Oyster Aquaculture: Estuarine Remediation via Site Suitability and Spatially Explicit Carrying Capacity Modeling in Virginia's Chesapeake Bay (2013).
21. *Angeleta Oveta Wilson:* A Participatory GIS approach to *unearthing* Manchester's Cultural Heritage '*gold mine'* (2013).
22. *Ola Svensson:* Visibility and Tholos Tombs in the Messenian Landscape: A Comparative Case Study of the Pylian Hinterlands and the Soulima Valley (2013).
23. *Monika Ogden:* Land use impact on water quality in two river systems in South Africa (2013).
24. *Stefan Rova:* A GIS based approach assessing phosphorus load impact on Lake Flaten in Salem, Sweden (2013).
25. *Yann Buhot:* Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitat diversity? (2013).
26. *Christina Fotiou:* Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence (2014).
27. *Inese Linuza:* Accuracy Assessment in Glacier Change Analysis (2014).
28. *Agnieszka Griffin:* Domestic energy consumption and social living standards: a GIS analysis within the Greater London Authority area (2014).
29. *Brynja Guðmundsdóttir:* Detection of potential arable land with remote sensing and GIS - A Case Study for Kjósarhreppur (2014).
30. *Oleksandr Nekrasov:* Processing of MODIS Vegetation Indices for analysis of agricultural droughts in the southern Ukraine between the years 2000-2012 (2014).
31. *Sarah Tressel:* Recommendations for a polar Earth science portal in the context of Arctic Spatial Data Infrastructure (2014).
32. *Caroline Gevaert:* Combining Hyperspectral UAV and Multispectral Formosat-2 Imagery for Precision Agriculture Applications (2014).
33. *Salem Jamal-Uddeen:* Using GeoTools to implement the multi-criteria evaluation analysis - weighted linear combination model (2014).
34. *Samanah Seyedi-Shandiz:* Schematic representation of geographical railway network at the Swedish Transport Administration (2014).
35. *Kazi Masel Ullah:* Urban Land-use planning using Geographical Information System and analytical hierarchy process: case study Dhaka City (2014).
36. *Alexia Chang-Wailing Spitteler:* Development of a web application based on MCDA and GIS for the decision support of river and floodplain rehabilitation projects (2014).
37. *Alessandro De Martino:* Geographic accessibility analysis and evaluation of potential changes to the public transportation system in the City of Milan (2014).
38. *Alireza Mollasalehi:* GIS Based Modelling for Fuel Reduction Using Controlled Burn in Australia. Case Study: Logan City, QLD (2015).
39. *Negin A. Sanati:* Chronic Kidney Disease Mortality in Costa Rica; Geographical Distribution, Spatial Analysis and Non-traditional Risk Factors (2015).
40. *Karen McIntyre:* Benthic mapping of the Bluefields Bay fish sanctuary, Jamaica (2015).
41. *Kees van Duijvendijk:* Feasibility of a low-cost weather sensor network for agricultural purposes: A preliminary assessment (2015).

42. *Sebastian Andersson Hylander:* Evaluation of cultural ecosystem services using GIS (2015).
43. *Deborah Bowyer:* Measuring Urban Growth, Urban Form and Accessibility as Indicators of Urban Sprawl in Hamilton, New Zealand (2015).
44. *Stefan Arvidsson:* Relationship between tree species composition and phenology extracted from satellite data in Swedish forests (2015).
45. *Damián Giménez Cruz*: GIS-based optimal localisation of beekeeping in rural Kenya (2016).
46. *Alejandra Narváez Vallejo:* Can the introduction of the topographic indices in LPJ-GUESS improve the spatial representation of environmental variables? (2016).
47. *Anna Lundgren:* Development of a method for mapping the highest coastline in Sweden using breaklines extracted from high resolution digital elevation models (2016).
48. *Oluwatomi Esther Adejoro:* Does location also matter? A spatial analysis of social achievements of young South Australians (2016).
49. *Hristo Dobrev Tomov:* Automated temporal NDVI analysis over the Middle East for the period 1982 - 2010 (2016).
50. *Vincent Muller:* Impact of Security Context on Mobile Clinic Activities A GIS Multi Criteria Evaluation based on an MSF Humanitarian Mission in Cameroon (2016).
51. *Gezahagn Negash Seboka:* Spatial Assessment of NDVI as an Indicator of Desertification in Ethiopia using Remote Sensing and GIS (2016).
52. *Holly Buhler:* Evaluation of Interfacility Medical Transport Journey Times in Southeastern British Columbia. (2016).
53. *Lars Ole Grottenberg*: Assessing the ability to share spatial data between emergency management organisations in the High North (2016).
54. *Sean Grant:* The Right Tree in the Right Place: Using GIS to Maximize the Net Benefits from Urban Forests (2016).
55. *Irshad Jamal:* Multi-Criteria GIS Analysis for School Site Selection in Gorno-Badakhshan Autonomous Oblast, Tajikistan (2016).
56. *Fulgencio Sanmartín:* Wisdom-volkano: A novel tool based on open GIS and time-series visualization to analyse and share volcanic data (2016).
57. *Nezha Acil:* Remote sensing-based monitoring of snow cover dynamics and its influence on vegetation growth in the Middle Atlas Mountains (2016).
58. *Julia Hjalmarsson:* A Weighty Issue: Estimation of Fire Size with Geographically Weighted Logistic Regression (2016).
59. *Mathewos Tamiru Amato:* Using multi-criteria evaluation and GIS for chronic food and nutrition insecurity indicators analysis in Ethiopia (2016).
60. *Karim Alaa El Din Mohamed Soliman El Attar:* Bicycling Suitability in Downtown, Cairo, Egypt (2016).
61. *Gilbert Akol Echelai:* Asset Management: Integrating GIS as a Decision Support Tool in Meter Management in National Water and Sewerage Corporation (2016).
62. *Terje Slinning:* Analytic comparison of multibeam echo soundings (2016).
63. *Gréta Hlín Sveinsdóttir:* GIS-based MCDA for decision support: A framework for wind farm siting in Iceland (2017).
64. *Jonas Sjögren:* Consequences of a flood in Kristianstad, Sweden: A GIS-based analysis of impacts on important societal functions (2017).

65. *Nadine Raska:* 3D geologic subsurface modelling within the Mackenzie Plain, Northwest Territories, Canada (2017).
66. *Panagiotis Symeonidis*: Study of spatial and temporal variation of atmospheric optical parameters and their relation with PM 2.5 concentration over Europe using GIS technologies (2017).
67. *Michaela Bobeck:* A GIS-based Multi-Criteria Decision Analysis of Wind Farm Site Suitability in New South Wales, Australia, from a Sustainable Development Perspective (2017).
68. *Raghdaa Eissa*: Developing a GIS Model for the Assessment of Outdoor Recreational Facilities in New Cities Case Study: Tenth of Ramadan City, Egypt (2017).
69. *Zahra Khais Shahid*: Biofuel plantations and isoprene emissions in Svea and Götaland (2017).
70. *Mirza Amir Liaquat Baig*: Using geographical information systems in epidemiology: Mapping and analyzing occurrence of diarrhea in urban - residential area of Islamabad, Pakistan (2017).
71. *Joakim Jörwall*: Quantitative model of Present and Future well-being in the EU-28: A spatial Multi-Criteria Evaluation of socioeconomic and climatic comfort factors (2017).
72. *Elin Haettner*: Energy Poverty in the Dublin Region: Modelling Geographies of Risk (2017).
73. *Harry Eriksson*: Geochemistry of stream plants and its statistical relations to soil- and bedrock geology, slope directions and till geochemistry. A GIS-analysis of small catchments in northern Sweden (2017).
74. *Daniel Gardevärn:* PPGIS and Public meetings – An evaluation of public participation methods for urban planning (2017).
75. *Kim Friberg:* Sensitivity Analysis and Calibration of Multi Energy Balance Land Surface Model Parameters (2017).
76. *Viktor Svanerud:* Taking the bus to the park? A study of accessibility to green areas in Gothenburg through different modes of transport (2017).
77. *Lisa-Gaye Greene*: Deadly Designs: The Impact of Road Design on Road Crash Patterns along Jamaica's North Coast Highway (2017).
78. *Katarina Jemec Parker*: Spatial and temporal analysis of fecal indicator bacteria concentrations in beach water in San Diego, California (2017).
79. *Angela Kabiru*: An Exploratory Study of Middle Stone Age and Later Stone Age Site Locations in Kenya's Central Rift Valley Using Landscape Analysis: A GIS Approach (2017).
80. *Kristean Björkmann*: Subjective Well-Being and Environment: A GIS-Based Analysis (2018).
81. *Williams Erhunmonmen Ojo*: Measuring spatial accessibility to healthcare for people living with HIV-AIDS in southern Nigeria (2018).
82. *Daniel Assefa*: Developing Data Extraction and Dynamic Data Visualization (Styling) Modules for Web GIS Risk Assessment System (WGRAS). (2018).
83. *Adela Nistora*: Inundation scenarios in a changing climate: assessing potential impacts of sea-level rise on the coast of South-East England (2018).
84. *Marc Seliger*: Thirsty landscapes - Investigating growing irrigation water consumption and potential conservation measures within Utah's largest master-planned community: Daybreak (2018).
85. *Luka Jovičić*: Spatial Data Harmonisation in Regional Context in Accordance with INSPIRE Implementing Rules (2018).

86. *Christina Kourdounouli*: Analysis of Urban Ecosystem Condition Indicators for the Large Urban Zones and City Cores in EU (2018).
87. *Jeremy Azzopardi*: Effect of distance measures and feature representations on distance-based accessibility measures (2018).
88. *Patrick Kabatha*: An open source web GIS tool for analysis and visualization of elephant GPS telemetry data, alongside environmental and anthropogenic variables (2018).
89. *Richard Alphonce Giliba*: Effects of Climate Change on Potential Geographical Distribution of Prunus africana (African cherry) in the Eastern Arc Mountain Forests of Tanzania (2018).
90. *Eiður Kristinn Eiðsson*: Transformation and linking of authoritative multi-scale geodata for the Semantic Web: A case study of Swedish national building data sets (2018).
91. *Niamh Harty*: HOP!: a PGIS and citizen science approach to monitoring the condition of upland paths (2018).
92. *José Estuardo Jara Alvear*: Solar photovoltaic potential to complement hydropower in Ecuador: A GIS-based framework of analysis (2018).
93. *Brendan O'Neill*: Multicriteria Site Suitability for Algal Biofuel Production Facilities (2018).
94. *Roman Spataru*: Spatial-temporal GIS analysis in public health – a case study of polio disease (2018).
95. *Alicja Miodońska*: Assessing evolution of ice caps in Suðurland, Iceland, in years 1986 - 2014, using multispectral satellite imagery (2019).
96. *Dennis Lindell Schettini*: A Spatial Analysis of Homicide Crime's Distribution and Association with Deprivation in Stockholm Between 2010-2017 (2019).
97. *Damiano Vesentini*: The Po Delta Biosphere Reserve: Management challenges and priorities deriving from anthropogenic pressure and sea level rise (2019).
98. *Emilie Arnesten*: Impacts of future sea level rise and high water on roads, railways and environmental objects: a GIS analysis of the potential effects of increasing sea levels and highest projected high water in Scania, Sweden (2019).
99. *Syed Muhammad Amir Raza*: Comparison of geospatial support in RDF stores: Evaluation for ICOS Carbon Portal metadata (2019).
100. *Hemin Tofiq*: Investigating the accuracy of Digital Elevation Models from UAV images in areas with low contrast: A sandy beach as a case study (2019).
101. *Evangelos Vafeiadis*: Exploring the distribution of accessibility by public transport using spatial analysis. A case study for retail concentrations and public hospitals in Athens (2019).
102. *Milan Sekulic*: Multi-Criteria GIS modelling for optimal alignment of roadway by-passes in the Tlokweng Planning Area, Botswana (2019).
103. *Ingrid Piirisaar*: A multi-criteria GIS analysis for siting of utility-scale photovoltaic solar plants in county Kilkenny, Ireland (2019).
104. *Nigel Fox*: Plant phenology and climate change: possible effect on the onset of various wild plant species' first flowering day in the UK (2019).
105. *Gunnar Hesch*: Linking conflict events and cropland development in Afghanistan, 2001 to 2011, using MODIS land cover data and Uppsala Conflict Data Programme (2019).
106. *Elijah Njoku*: Analysis of spatial-temporal pattern of Land Surface Temperature (LST) due to NDVI and elevation in Ilorin, Nigeria (2019).

107. *Katalin Bunyevácz*: Development of a GIS methodology to evaluate informal urban green areas for inclusion in a community governance program (2019).
108. *Paul dos Santos*: Automating synthetic trip data generation for an agent-based simulation of urban mobility (2019).
109. *Robert O' Dwyer*: Land cover changes in Southern Sweden from the mid-Holocene to present day: Insights for ecosystem service assessments (2019).
110. *Daniel Klingmyr*: Global scale patterns and trends in tropospheric NO2 concentrations (2019).
111. *Marwa Farouk Elkabbany*: Sea Level Rise Vulnerability Assessment for Abu Dhabi, United Arab Emirates (2019).
112. *Jip Jan van Zoonen*: Aspects of Error Quantification and Evaluation in Digital Elevation Models for Glacier Surfaces (2020).
113. *Georgios Efthymiou*: The use of bicycles in a mid-sized city – benefits and obstacles identified using a questionnaire and GIS (2020).
114. *Haruna Olayiwola Jimoh*: Assessment of Urban Sprawl in MOWE/IBAFO Axis of Ogun State using GIS Capabilities (2020).
115. *Nikolaos Barmpas Zachariadis*: Development of an iOS, Augmented Reality for disaster management (2020).
116. *Ida Storm*: ICOS Atmospheric Stations: Spatial Characterization of CO2 Footprint Areas and Evaluating the Uncertainties of Modelled CO2 Concentrations (2020).
117. *Alon Zuta*: Evaluation of water stress mapping methods in vineyards using airborne thermal imaging (2020).
118. *Marcus Eriksson*: Evaluating structural landscape development in the municipality Upplands-Bro, using landscape metrics indices (2020).
119. *Ane Rahbek Vierø*: Connectivity for Cyclists? A Network Analysis of Copenhagen's Bike Lanes (2020).
120. *Cecilia Baggini*: Changes in habitat suitability for three declining Anatidae species in saltmarshes on the Mersey estuary, North-West England (2020).
121. *Bakrad Balabanian*: Transportation and Its Effect on Student Performance (2020).
122. *Ali Al Farid*: Knowledge and Data Driven Approaches for Hydrocarbon Microseepage Characterizations: An Application of Satellite Remote Sensing (2020).
123. *Bartlomiej Kolodziejczyk*: Distribution Modelling of Gene Drive-Modified Mosquitoes and Their Effects on Wild Populations (2020).
124. *Alexis Cazorla*: Decreasing organic nitrogen concentrations in European water bodies - links to organic carbon trends and land cover (2020).
125. *Kharid Mwakoba*: Remote sensing analysis of land cover/use conditions of community-based wildlife conservation areas in Tanzania (2021).
126. *Chinatsu Endo*: Remote Sensing Based Pre-Season Yellow Rust Early Warning in Oromia, Ethiopia (2021).
127. *Berit Mohr*: Using remote sensing and land abandonment as a proxy for long-term human out-migration. A Case Study: Al-Hassakeh Governorate, Syria (2021).
128. *Kanchana Nirmali Bandaranayake*: Considering future precipitation in delineation locations for water storage systems - Case study Sri Lanka (2021).

129. *Emma Bylund*: Dynamics of net primary production and food availability in the aftermath of the 2004 and 2007 desert locust outbreaks in Niger and Yemen (2021).
130. *Shawn Pace*: Urban infrastructure inundation risk from permanent sea-level rise scenarios in London (UK), Bangkok (Thailand) and Mumbai (India): A comparative analysis (2021).
131. *Oskar Evert Johansson*: The hydrodynamic impacts of Estuarine Oyster reefs, and the application of drone technology to this study (2021).
132. *Pritam Kumarsingh*: A Case Study to develop and test GIS/SDSS methods to assess the production capacity of a Cocoa Site in Trinidad and Tobago (2021).
133. *Muhammad Imran Khan*: Property Tax Mapping and Assessment using GIS (2021).
134. *Domna Kanari*: Mining geosocial data from Flickr to explore tourism patterns: The case study of Athens (2021).
135. *Mona Tykesson Klubien*: Livestock-MRSA in Danish pig farms (2021).
136. *Ove Njøten*: Comparing radar satellites. Use of Sentinel-1 leads to an increase in oil spill alerts in Norwegian waters (2021).
137. *Panagiotis Patrinos*: Change of heating fuel consumption patterns produced by the economic crisis in Greece (2021).
138. *Lukasz Langowski*: Assessing the suitability of using Sentinel-1A SAR multi-temporal imagery to detect fallow periods between rice crops (2021).
139. *Jonas Tillman*: Perception accuracy and user acceptance of legend designs for opacity data mapping in GIS (2022).
140. *Gabriela Olekszyk*: ALS (Airborne LIDAR) accuracy: Can potential low data quality of ground points be modelled/detected? Case study of 2016 LIDAR capture over Auckland, New Zealand (2022).
141. *Luke Aspland*: Weights of Evidence Predictive Modelling in Archaeology (2022).
142. *Luís Fareleira Gomes*: The influence of climate, population density, tree species and land cover on fire pattern in mainland Portugal (2022).
143. *Andreas Eriksson*: Mapping Fire Salamander (Salamandra salamandra) Habitat Suitability in Baden-Württemberg with Multi-Temporal Sentinel-1 and Sentinel-2 Imagery (2022).
144. *Lisbet Hougaard Baklid*: Geographical expansion rate of a brown bear population in Fennoscandia and the factors explaining the directional variations (2022).
145. *Victoria Persson*: Mussels in deep water with climate change: Spatial distribution of mussel (Mytilus galloprovincialis) growth offshore in the French Mediterranean with respect to climate change scenario RCP 8.5 Long Term and Integrated Multi-Trophic Aquaculture (IMTA) using Dynamic Energy Budget (DEB) modelling (2022).
146. *Benjamin Bernard Fabien Gérard Borgeais*: Implementing a multi-criteria GIS analysis and predictive modelling to locate Upper Palaeolithic decorated caves in the Périgord noir, France (2022).
147. *Bernat Dorado-Guerrero*: Assessing the impact of post-fire restoration interventions using spectral vegetation indices: A case study in El Bruc, Spain (2022).

148. *Ignatius Gabriel Aloysius Maria Perera*: The Influence of Natural Radon Occurrence on the Severity of the COVID-19 Pandemic in Germany: A Spatial Analysis (2022).
149. *Mark Overton*: An Analysis of Spatially-enabled Mobile Decision Support Systems in a Collaborative Decision-Making Environment (2022).
150. *Viggo Lunde*: Analysing methods for visualizing time-series datasets in open-source web mapping (2022).
151. *Johan Viscarra Hansson*: Distribution Analysis of Impatiens glandulifera in Kronoberg County and a Pest Risk Map for Alvesta Municipality (2022).
152. *Vincenzo Poppiti*: GIS and Tourism: Developing strategies for new touristic flows after the Covid-19 pandemic (2022).
153. *Henrik Hagelin*: Wildfire growth modelling in Sweden - A suitability assessment of available data (2023).
154. *Gabriel Romeo Ferriols Pavico*: Where there is road, there is fire (influence): An exploratory study on the influence of roads in the spatial patterns of Swedish wildfires of 2018 (2023).
155. *Colin Robert Potter*: Using a GIS to enable an economic, land use and energy output comparison between small wind powered turbines and large-scale wind farms: the case of Oslo, Norway (2023).
156. *Krystyna Muszel*: Impact of Sea Surface Temperature and Salinity on Phytoplankton blooms phenology in the North Sea (2023).
157. *Tobias Rydlinge*: Urban tree canopy mapping - an open source deep learning approach (2023).
158. *Albert Wellendorf*: Multi-scale Bark Beetle Predictions Using Machine Learning (2023).
159. *Manolis Papadakis*: Use of Satellite Remote Sensing for Detecting Archaeological Features: An Example from Ancient Corinth, Greece (2023).
160. *Konstantinos Sourlamtas*: Developing a Geographical Information System for a water and sewer network, for monitoring, identification and leak repair - Case study: Municipal Water Company of Naoussa, Greece (2023).
161. *Xiaoming Wang*: Identification of restoration hotspots in landscape-scale green infrastructure planning based on model-predicted connectivity forest (2023).
162. *Sarah Sienaert*: Usability of Sentinel-1 C-band VV and VH SAR data for the detection of flooded oil palm (2023).
163. *Katarina Ekeroot*: Uncovering the spatial relationships between Covid-19 vaccine coverage and local politics in Sweden (2023).
164. *Nikolaos Kouskoulis*: Exploring patterns in risk factors for bark beetle attack during outbreaks triggered by drought stress with harvester data on attacked trees: A case study in Southeastern Sweden (2023).
165. *Jonas Almén*: Geographic polarization and clustering of partisan voting: A local-level analysis of Stockholm Municipality (2023).
166. *Sara Sharon Jones*: Tree species impact on Forest Fire Spread Susceptibility in Sweden (2023).