Florian Zellweger
06-704-852

Structural diversity in Swiss mountain forests

An assessment of hazel grouse (*Bonasa bonasia*) habitat based on field data and laser remote sensing

GEO 511
Master thesis

Advisers

Dr. Kurt Bollmann (Swiss Federal Research Institute WSL)
Prof. Dr. Ross Purves (Faculty Representative)
Dr. Felix Morsdorf

GIScience, Department of Geography, University of Zurich - Irchel

Submitted 31st January 2012

Contact

Advisers

Dr. Kurt Bollmann

Biodiversity and Conservation Biology, Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland.

Prof. Dr. Ross Purves

GIScience, Department of Geography, University of Zurich - Irchel, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

Dr. Felix Morsdorf

Remote Sensing Laboratories, Department of Geography, University of Zurich - Irchel, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

Author

Florian Zellweger

Friesenbergstrasse 88, 8055 Zurich, Switzerland.
flo.zellweger@gmail.com

Summary

Habitat quality and biodiversity in forest ecosystems largely depends on structural diversity, which is difficult to measure in the field, especially across broad spatial scales. However, emerging remote sensing technologies such as light detection and ranging (LiDAR) provide alternative means to quantify structural diversity across relatively broad areas at a fine resolution. In order to quantify the structural diversity in Swiss mountain forests, the habitat requirements of hazel grouse (*Bonasa bonasia*), an indicator species of structurally rich forest stands, were assessed over three biogeographical regions in Switzerland.

Based on 1km^2 presence/absence data, a twofold habitat assessment was carried out, encompassing a sample based field survey and a LiDAR based, area-wide assessment of vertical and horizontal forest structures. Species distribution models were computed for both variable sets separately and in combination, using boosted regression trees, a statistical machine learning technique.

The combined model performed best, with over 55% explained deviance in the observed data, followed by the field (37%) and LiDAR model (34%). The discrimination power based on AUC values was outstanding for the combined model (0.977), and excellent for the field (0.892) and LiDAR models (0.884). Evaluation measures based on predictions to left-out data using cross-validation ranged from 18% to 23% explained deviance, with AUC values indicating a good discrimination power. The field survey revealed that hazel grouse favoured an evenly distributed, rich ground vegetation, optimally in combination with bilberry (*Vaccinium myrtillus*). The abundance of tall rowans (*Sorbus aucuparia*), basal branched trees and a high percentage of resource trees in the shrub layer were found to be further critical habitat elements. In addition to reflecting the current knowledge of hazel grouse habitat, the results comprehend variable thresholds that were derived at a relevant spatial scale (1km^2) for forest management.

LiDAR was powerful in detecting important structural features, whereby the horizontal forest structure explained more of the deviance than the vertical forest structure. The most influential LiDAR variable was a measure of canopy height heterogeneity. Apart from indicating a rich structure, it probably also served as a proxy of compositional aspects such as the abundance of light demanding resource trees and shrubs or of a well developed field layer.

This study showed that LiDAR provides adequate means to assess structural habitat elements area-wide, thus overcoming the difficulties associated with sample based field assessments. The methods presented here create new insights into the habitat connectivity, which is essential for the conservation of viable populations. In addition, the methods can further be developed to identify hotspots of forest structural richness, a matter of interest in the light of emerging attempts to conserve biodiversity in forests.

Zusammenfassung

Die Lebensraumqualität und Biodiversität in Waldökosystemen hängt wesentlich von der strukturellen Vielfalt des Waldes ab. Strukturelle Waldparameter sind jedoch relativ schwierig zu messen im Feld, vor allem wenn grossräumige Gebiete erfasst werden müssen. Neuere Fernerkundungsmethoden wie flugzeuggetragene Laserscanner (Light detection and ranging, LiDAR) bieten Möglichkeiten, um die strukturelle Vielfalt über relativ grossräumige Flächen mit hoher Präzision zu erfassen. Um die strukturelle Vielfalt in Schweizer Gebirgswäldern zu quantifizieren, wurden die Habitatansprüche des Haselhuhns (*Bonasa bonasia*) in drei biogeographischen Regionen der Schweiz untersucht. Das Haselhuhn wurde gewählt, weil es als Indikatorart für reichhaltig strukturierte Wälder angesehen wird, was sich in den Habitatansprüchen widerspiegelt.

Anhand von 1km^2 Präsenz/Absenz-Daten wurde eine zweifache Habitatanalyse durchgeführt. Einerseits lieferte eine stichprobenbasierte Felderhebung Daten über Struktur und Zusammensetzung des Waldes, andererseits wurden anhand von LiDAR-Daten Parameter der vertikalen und horizontalen Waldstruktur flächendeckend abgeleitet. Mittels 'boosted regression trees', einer statistischen Technik des maschinellen Lernens, wurden Artverbreitungsmodelle für beide Datensätze separat, sowie für deren Kombination gerechnet.

Das kombinierte Modell erklärte den grössten Anteil der Varianz der Beobachtungsdaten (55%), gefolgt vom feldbasierten Modell (37%) und dem LiDAR Modell (34%). Die Modelldiskriminierungskraft basierend auf den AUC-Werten waren hervorragend im kombinierten Modell (0.977) und sehr gut im feldbasierten (0.892) und LiDAR Modell (0.884). Die Evaluationskriterien anhand von kreuzvalidierten Vorhersagen für zurückbehaltene Testdaten fielen etwas tiefer aus. Sie erreichten Werte von 18% bis 23% erklärter Abweichung, sowie AUC-Werte von rund 0.8, was auf eine gute Diskriminierung hindeutet. Die Felderhebungen haben gezeigt, dass das Haselhuhn eine gleichmässig verteilte, reichhaltige Bodenvegetation bevorzugt, in der vorzugsweise etwas Heidelbeere (*Vaccinium myrtillus*) zu finden ist. Das Vorkommen von grossen Vogelbeeren (*Sorbus aucuparia*), tiefbeasteten Bäumen und ein hoher Anteil an Nahrungsbäumen und -sträuchern im Unterwuchs waren weitere Habitatmerkmale. Diese Ergebnisse stimmen weitgehend mit den Resultaten anderer Haselhuhn-Habitatanalysen überein und erlaubten es zusätzlich, Schwellenwerte der Variablen auf einer für die Forstwirtschaft angemessenen räumlichen Skala (1km^2) abzuleiten.

Die LiDAR basierten Parameter erfassten wichtige strukturelle Merkmale, wobei die horizontale Struktur einen grösseren Teil der erklärten Abweichung ausmachte als die vertikale Struktur. Der einflussreichste LiDAR-Parameter war ein Mass für die Heterogenität der Kronendachhöhe. Dies deutet darauf hin, dass dieser Parameter nicht nur ein Indikator einer reichhaltigen Struktur ist, sondern indirekt auch auf gewisse Aspekte der Pflan-

zenzusammensetzung hindeutet, wie beispielsweise das Vorkommen von lichtbedürftigen Nahrungsbäumen und -sträuchern, sowie einer gut ausgebildeten Bodenvegetation.

Diese Studie zeigte, dass LiDAR eine geeignete Technik ist zur grossräumigen Erfassung von strukturellen Habitateigenschaften, wobei die Schwierigkeiten mit Stichprobenbasierten Feldaufnahmen überwunden werden konnten. Die hier entwickelten Methoden erlauben es künftig, die strukturelle und funktionelle Lebensraumkonnektivität flächendeckend zu erheben und in der Waldentwicklungsplanung zu nutzen, was wiederum für den Schutz von Populationen gefährdeter Arten von Bedeutung ist. Ferner können auch LiDAR basierte Parameter für die Identifikation von hotspots struktureller Vielfalt generiert werden, die im Rahmen künftiger Bestrebungen zum Erhalt der Biodiversität im Wald von grossem Interesse sind.

Contents

1 Introduction **1**

1.1 Framework of this thesis . . . 2
1.2 State of the Art . . . 2
1.2.1 Swiss forests, hazel grouse habitat and its conservation . . . 2
1.2.2 Forest structure and LiDAR . . . 3
1.3 Research gap . . . 4
1.4 Objectives and research questions . . . 4

2 Methods **6**

2.1 Study area . . . 6
2.2 Study design and species data . . . 6
2.3 Habitat assessment . . . 11
2.3.1 Field survey . . . 11
2.3.2 LiDAR based habitat assessment . . . 13
2.4 Statistical/Machine learning analysis . . . 23
2.4.1 Boosted regression trees . . . 23
2.4.2 Model fitting and evaluation . . . 24

3 Results and interpretation **27**

3.1 Model performance . . . 27
3.2 Variable contribution . . . 28
3.2.1 Field model . . . 28
3.2.2 LiDAR model . . . 29
3.2.3 Combined model . . . 29
3.3 Threshold effects . . . 29
3.3.1 Field model . . . 29
3.3.2 LiDAR model . . . 31
3.3.3 Combined model . . . 31
3.4 Variable interactions . . . 33
3.5 Altitude and regional aspects . . . 33

4 Discussion **35**

4.1 Study design . . . 35
4.2 Model comparison and assessment . . . 36
4.3 Hazel grouse habitat characteristics in mountain forests . . . 37
4.3.1 Structural diversity . . . 39
4.4 LiDAR based habitat assessments of forest species . . . 41

5 Conclusions **44**
5.1 Achievements 44
5.2 Insights and outlook 44

A Statistical overview of field and LiDAR derived habitat variables **56**

B MATLAB code for processing LiDAR data **58**

C Personal declaration **60**

List of Figures

1 Conceptual overview of methodology 7
2 Study area 8
3 Sampling design including 16 sampling plots per km^2 10
4 Field plot sampling design 13
5 Three-dimensional representation of a LiDAR raw data point cloud 14
6 Flow-chart for processing LiDAR data 17
7 Orthogonal projection of terrain corrected vegetation heights for a $1km^2$ sample cell 18
8 Classified Canopy Height Models (CHM) showing a presence and a pseudo absence km^2 cell 19
9 Multi-band raster of terrain corrected vegetation point heights 20
10 Simplification of a BRT model based on variable exclusion and change in predictive deviance 25
11 Partial dependence plots for variables retained by the final BRT field and LiDAR model 30
12 Partial dependence plots for 9 most influential variables retained by the final BRT combined model 32
13 Dense ground vegetation with patches of dwarf-shrubs 38
14 Fructifying rowan trees and other deciduous softwoods growing in a former Lothar windthrow area. 39
15 Extract of classified canopy height model showing continuous transitions from forest to open land 41
16 Vertical distribution of vegetation point heights in a multi-layered and single-layered forest stand 42

List of Tables

1 Field survey sample size per study region 10
2 Field variables used to quantify habitat characteristics in a sampling plot . 12
3 Key characteristics of Swisstopo LiDAR data 14
4 Sample size of LiDAR based habitat assessment per study region. 15
5 Predictor variables derived from the vegetation point cloud 21
6 Predictor variables based on FRAGSTATS metrics 22
7 Optimal settings, performance and AUC values for best performing BRT models 27
8 Predictor variables contributions to the total deviance explained for the field, LiDAR and combined model 28
9 Variable interactions in the combined model 33
10 Optimal settings, performance and AUC values for best performing BRT models including the variable altitude, and excluding the Central Eastern Alps 34
11 Statistical overview of field and LiDAR derived habitat variables 57

1 Introduction

Forest ecosystems once covered most of Europe and are still widespread, but only few forests have been spared from human intervention [56]. In Central Europe, for example, only 0.2% of deciduous forests remain in a relatively natural state [34]. The structure of such pristine forests is driven by the prevailing natural disturbance regime that allows late-successional stages to develop, resulting in a structurally diverse mosaic of different development stages [4, 23]. Structural diversity or richness, in turn, influences the number of potential niches that is provided by a forest plant community and is therefore an important driver of wildlife habitat and biodiversity in general [4]. In managed forest, however, structure and therefore the habitat quality and biodiversity is to a large extent determined by the management type and intensity [76].

Forest management has altered key attributes of most European forest ecosystems, such as their structure and composition. Today, this is reflected in a general reduction in stand age resulting from reduced rotation lengths, and in more homogeneous tree compositions through the promotion of commercially preferred tree species [56]. This entailed a major reduction of old-growths stands and areas of natural rejuvenation with pioneer species, which is of conservational concern because many species depend on the structural diversity provided by such stands or development stages [4]. The structural diversity of a forest stand compromises individual elements as well as their spatial arrangement. Structural elements include the tree species, size and condition as well as standing dead trees and coarse woody debris on the forest floor. The spatial arrangement of such elements include the vertical distribution of branches and foliage, and the horizontal distribution of trees (e.g. random, dispersed, or aggregated) and other structures (e.g. gaps) [23].

In order to improve conservation attempts, knowledge about the spatial distribution of forest structural diversity is essential. Field assessments can help to provide this knowledge, but such surveys are time and cost intensive and their spatial extent is often restricted by the sampling size and design, which limit the scales of analysis. In addition, field assessments can even be dangerous in difficult terrain or remote areas. However, finding ways to quantify patterns of variability across multiple scales is a fundamental task in ecology and conservation biology [61]. Remote sensing can provide one potential route to meet these requirements and is nowadays widely used for habitat and biodiversity assessments [84]. Light detection and ranging (LiDAR), in particular, is a relatively new active remote sensing technique that can quantify three-dimensional structures in forests and other ecosystems over large areas at a high level of detail [59]. This makes it a promising tool for analysing area-wide structural patterns and species-habitat relationships [86, 5].

1.1 Framework of this thesis

In 2008, the Federal Office for the Environment (FOEN) and the Swiss Federal Research Institute for Forest, Snow and Landscape Research (WSL) launched a broad scientific program with aiming to assess the impact of climate change on Swiss forest ecosystems. The research project with the title *"Predicted effects of climate change on indicator species of structural and biological diversity in mountain forests: towards adaptive forest management in the face of environmental uncertainty"* is associated with this program and provided the framework of this thesis. In a first stage, the project involves an assessment of habitat characteristics for four forest bird species: Capercaillie (*Tetrao urogallus*), three-toed woodpecker (*Picoides tridactylus*), pygmy owl (*Glaucidium passerinum*) and hazel grouse (*Bonasa bonasia*). Among other reasons, the choice of species was based on the fact that they are all specialised to different structural elements and vegetation successional stages, and can thus serve as indicators of overall structural diversity [77, 9]. Each of these species was assigned to a Master student (four in total), who individually performed a habitat analysis at the local scale. Hence, the choice of hazel grouse as target species for this study.

1.2 State of the Art

1.2.1 Swiss forests, hazel grouse habitat and its conservation

During the past 150 years, the forested area in Switzerland has gradually increased and now covers 31% of the land [12]. Mountain forest management objectives in the late 19th century until about 1970 mainly aimed at increasing the forested area and standing volume as well as at enhancing the protective effect of forests [80]. During the same period, the switch to fossil fuels reduced pressure on forests and increasing wealth as well as falling timber prices resulted in reduced harvesting interventions [90]. This has led to a trend towards single-layered, high stands with a deficit in structural richness and heterogeneity [80, 8]. Compared to natural forests, this deficit is reflected by a scarce occurrence of old-growth forests and young seral stages with pioneer tree species [8].

Hazel grouse has suffered from these changes in forest management because it depends on structurally rich stands with a sufficient availability of light demanding resource trees or shrubs of the genera *Sorbus, Betula, Salix, Corylus, Crataegus, Alnus, Populus, Rosa* and *Rubus* [7, 6, 27]. Habitat fragmentation and forest succession in the lowland of Switzerland restricted hazel grouse populations to the mountain forests, where relevant habitat characteristics are still preserved by small-scale changes in site conditions, local disturbances (e.g. windthrows, snow movements, snow breakage, insects, diseases) and forest management practices in favour of resource trees and shrubs [79]. Forests in early seral stages with

pioneer tree species, which provide sufficient cover and resources, are optimal hazel grouse habitats, especially in wintertime, when other potential food sources in the field layer are not accessible due to snow cover [78, 79, 6]. Patches of natural rejuvenation embedded in old-growths forests or in the vicinity of brooks, mires or even forest roads also provide suitable habitats [70, 6].

Hazel grouse is a cryptic[1] habitat specialist whose present distribution in Central Europe is largely restricted to the Alpine region [51, 6]. Switzerland has a special international responsibility for hazel grouse because the remaining populations in the Swiss Alps account for a substantial part of the populations left in Central Europe [46]. Today, hazel grouse is classified as vulnerable in the red list of breeding bird species of Switzerland [47] and considered a priority species for species action plans [10].

Over the course of the last decades, multifunctional silviculture in Swiss mountain forests has started to recognise the role of natural processes (e.g. natural regeneration) and dynamics after disturbances (e.g. windstorms Vivian, 1990 and Lothar, 1999) [80]. Simultaneously, agricultural land use has changed and resulted in less intensive farming of alpine pastures. This in turn led to an increase of forest area, which is characterised by ingrowth of natural rejuvenation and pioneer trees and shrubs [25]. These environmental changes based on forest dynamics influence the structural diversity of mountain forests and are thus of significance not only for hazel grouse but also for a wide variety of other species [81]. Considering that hazel grouse is an indicator species of early successional stages and structurally rich forests, an assessment of its habitat requires adequate methods that are able to quantify these environmental changes.

1.2.2 Forest structure and LiDAR

The ability of LiDAR remote sensing to measure three-dimensional structures as well as biophysical parameters makes it an interesting tool for ecosystem studies [59, 93, 68]. LiDAR derived forest inventory data, for example, can be retrieved at a tree and stand level in more detail compared to other remote sensing methods [43]. Several characteristics such as mean tree height, stem number, basal area or timber volume were successfully derived on georeferenced sample plots and the estimates predicted for entire forest stands [73]. The results confirm that the applied procedure performs better than traditional inventory methods based on field measurements or aerial photo interpretation.

While LiDAR data is now operationally used in commercial, stand-based forest inventories [74], there is also a growing number of ecological applications [87]. Although reviewed by [86] and [5], studies regarding the use of LiDAR in biodiversity and habitat

[1]The word cryptic reflects the difficulty to observe this species due to its inconspicuous behaviour and camouflaged appearance

assessments are, however, still relatively rare. A comparison of LiDAR-derived variables of forest structure with field measurements and other remotely sensed data in a montane ecosystem in California revealed the potential of LiDAR variables in terms of mapping forest structure for wildlife habitat analysis [41, 42]. LiDAR metrics of the canopy vertical structure were successfully identified as predictors of bird species richness in a eastern temperate forest in the USA [29] and used to predict breeding habitat of the Black-throated Blue Warbler (*Dendroica caerulescens*) [28]. Similarly, LiDAR-derived vegetation height, density or volume metrics were found to correlate with avian species diversity and density in South Dakota, USA [14] and were used to model forest songbird species richness in Wisconsin, USA [60]. In Europe, LiDAR based canopy heights and percentage of canopy cover correlated well with bird distributions in lowland British conditions [38] and were used to predict the habitat quality of Great Tits (*Parus major*) [37]. In the Swiss Alps, habitat models for capercaillie (*Tetrao urogallus)* could be improved with structural variables from LiDAR data [32]. Further, LiDAR variables were superior to variable sets based on aerial photographs and field measurements in terms of modelling the potential abundance and assemblages of forest passerines in a montane forest in southeast Germany [72].

1.3 Research gap

Existing studies incorporating LiDAR data and habitat assessments are generally based on point observations of the target species or species groups and involve local scale analysis across a relatively small spatial extent (e.g. forest reserves or national parks). However, a particular strength of LiDAR is that the data is recorded across relatively broad areas at a fine resolution [59], thus allowing for analysis on different spatial scales. Considering the above cited literature, much effort has been devoted to quantifying the vertical forest structure, whereby horizontal structural patterns such as the arrangement of trees or forest patches, or ecotones like forest - non-forest transitions have received less attention. This work tries to bridge these gaps by developing adequate methods to quantify both, the vertical and horizontal structure of forests, applying area based species data that allow for an expansion of the study area over three biogeographic regions to the landscape level. Thus, it was anticipated to include much of the variability of hazel grouse habitat by broadly covering Central European mountain forest conditions as represented by three mountain regions in Switzerland.

1.4 Objectives and research questions

The aim of this study was to assess the habitat requirements of hazel grouse at the scale of 1km^2 across a broad range of mountain forest conditions in Central Europe. A field survey was designed to provide habitat data based on plot samples of forest structure and

composition. Further, the potential of LiDAR data for modelling habitat characteristics area-wide were explored by developing new methods to assess the regional forest structural diversity. In order to achieve these objectives, the following research questions were asked:

- How does species presence relate to forest composition and structure variables at the scale of 1km^2; what variable thresholds can be derived?
- What LiDAR derived variables of vertical and horizontal forest structures can be used to predict hazel grouse habitat suitability?
- How do field and LiDAR based assessments as well as the combination thereof compare to each other?

2 Methods

The methodological approach involves a twofold habitat assessment based on the study design and species data of the research project mentioned in 1.1. On the one hand, plot based field data samples of local habitat characteristics in terms of forest composition and structure were collected and aggregated for each sampling unit of 1km^2. On the other hand, area-wide available LiDAR data was processed and analysed in order to derive forest structural variables reflecting important habitat elements. The vertical forest structure was analysed based on variables derived from LiDAR point cloud statistics. The horizontal forest structure was assessed based on FRAGSTATS metrics, which allow for the analysis of spatial patterns and landscape structures. Both variable sets as well as a combination thereof were integrated into species distribution models, using boosted regression trees (BRT), a statistical machine learning technique [19]. An overview of the methodology is provided in Figure 1.

2.1 Study area

The study area consists of three biogeographic regions in Switzerland: the Swiss Jura (1), the Northern Prealps (2) and the Eastern Central Alps (3) (Figure 2). The biogeographical classification is based on distribution patterns of the flora and fauna and reflects numerous biotic and abiotic factors such as dispersal ranges of species before and after glacial periods, interspecific competition or geological and climatic conditions [30]. The three regions represent an ecological stratification with respect to mountain forest ecosystems as study objects. In the Swiss Jura, elevations do not reach the tree line and forests would thus naturally form continuous areas uninterrupted by alpine grasslands or mountain peaks. The altitude of studied forest areas in the Swiss Jura ranged from 708 to 1519 m.a.s.l. In the Northern Prealps, montane and subalpine forests have their upper limit at about 1800 m.a.s.l., and mountains with elevations of up to over 3000 m.a.s.l. occasionally punctuate the potential natural forest dispersion cover. In this region, the altitude of the studied forest areas ranges from 786 to 1683 m.a.s.l. The Eastern Central Alps are characterised by high relief over short distances, with mountains reaching elevations of up to 4000 m.a.s.l. The forests in this region occur in belts around mountain ranges, and the tree line lies at around 2300 m.a.s.l. Consequently, the studied forest areas in the Eastern Central Alps have the widest altitudinal range of 805 to 2020 m.a.s.l.

2.2 Study design and species data

As a reference grid for all sampling areas, the Swiss coordinate system with a cell size of 1 x 1km was used. Hazel grouse presence data were derived from the database at the Swiss

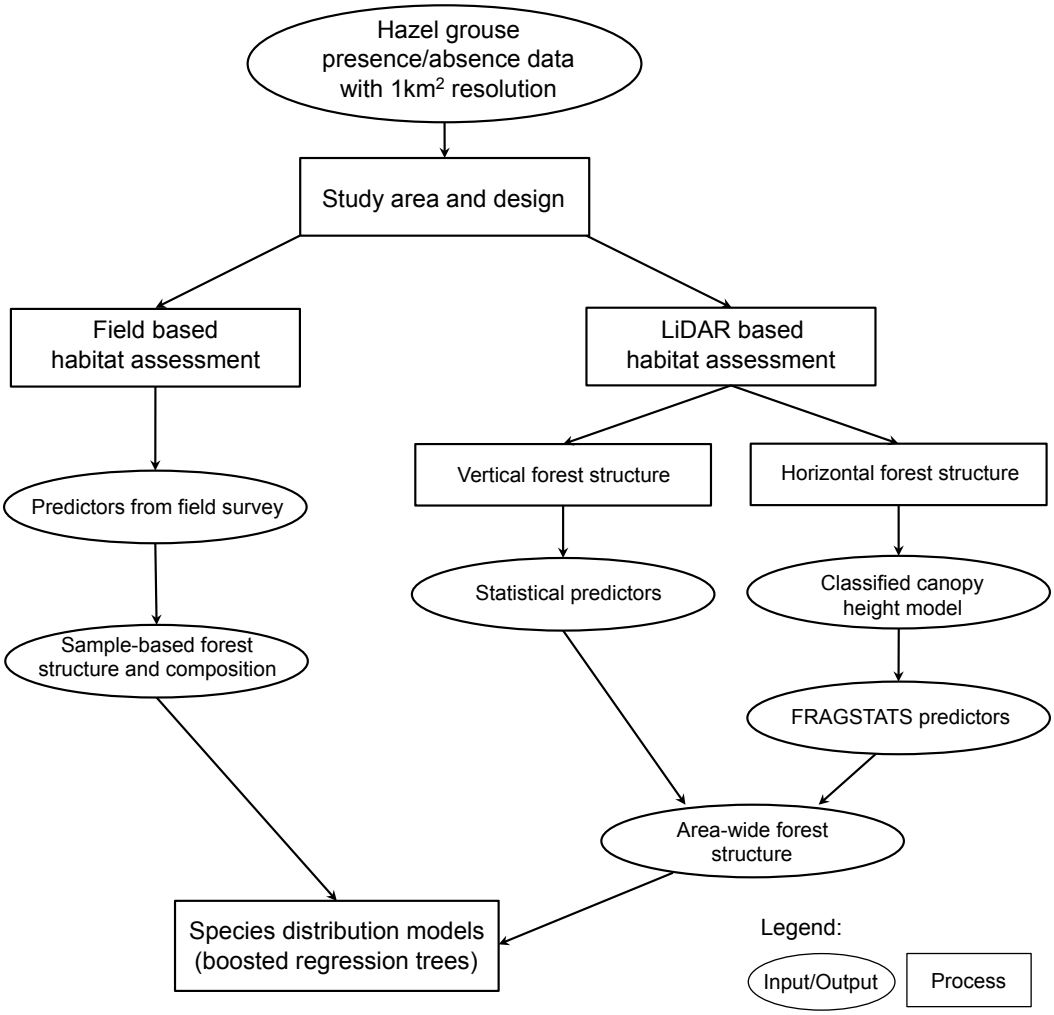


Figure 1: Conceptual overview of methodology.

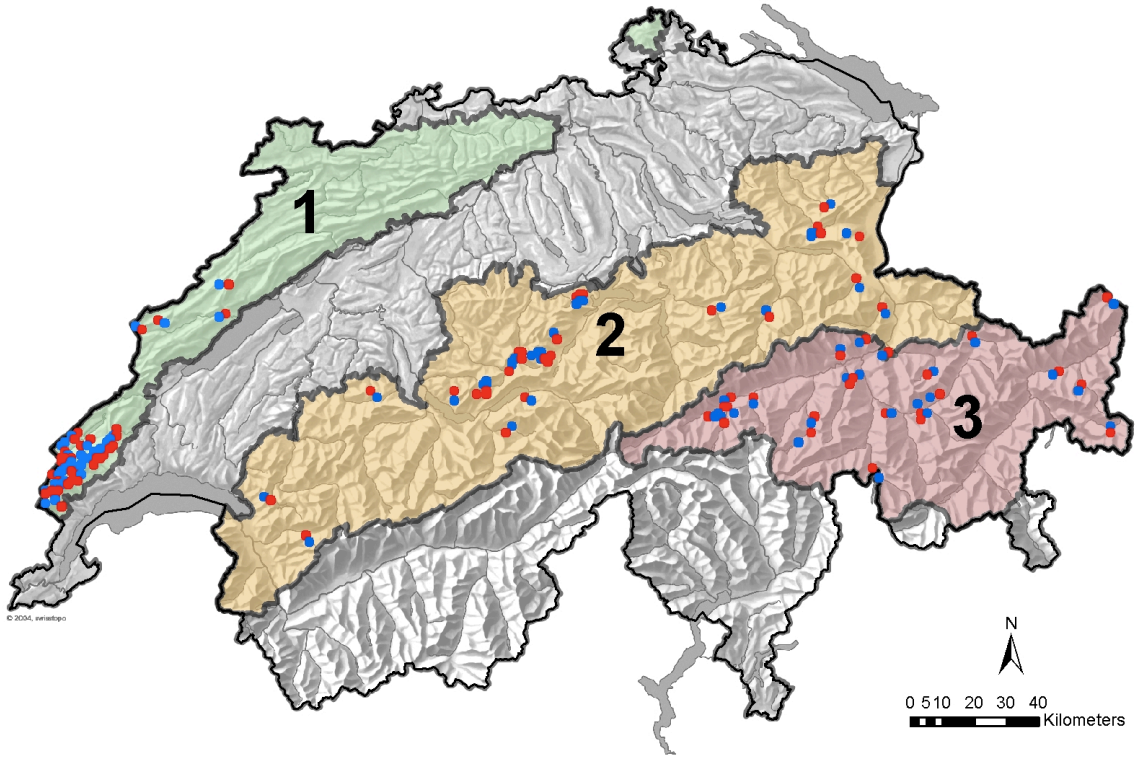


Figure 2: Study area compromising three biogeographic mountain regions in Switzerland. Jura (1), Northern Prealps (2), Central Eastern Alps (3). Coloured squares indicate presence (blue) and pseudo absence (red) cells. Regions delineated according to [30].

Ornithological Institute and the Grouse database of the Swiss Federal Research Institute WSL. These databases consist of field based species evidences which were assigned to their corresponding km^2 cell delineated by the Swiss coordinate system. Thus, the spatial resolution of the species data is $1km^2$. This corresponds approximately to the home range of hazel grouse, which is estimated to be around 50 to 70 hectares [27]. Presence cells were selected along an altitudinal gradient and allocated such that the distribution range of hazel grouse in each study region was represented. The cell selection was restricted to observations made within the past five years on three or more occasions. In addition, priority was given to cells with sympatric occurrences of the project's target species (see 1.1).

Pseudo absence cells were defined as having more than 50% of the cell area covered by forest (delineated by Swisstopo in Vector25 [75]) with no evidence of one of the target species within the last ten years. The term pseudo absence is used to indicate the uncertainty of whether the species actually uses these areas or not, and to indicate bias in recording presence data, since all that is certain is that *no evidence* has been *reported* within the last ten years. This may be because the habitat is unsuitable or simply because it is not frequently visited by people looking for evidence [45]. Because of the underlying idea that the pseudo absence cells could potentially be used but are unused due to suboptimal habitat characteristics, pairs of presence and pseudo absence cells were located in the same valley within dispersal distance of the target species. A paired multi-criteria model was applied to allocate the pseudo absence cells: starting from a presence cell being part of the Swiss coordinate system grid, a first order square of grid cells adjacent to the presence cell was designated, followed by a second, third and fourth order square. Considering the two above mentioned criteria forest cover and no target species evidence, a pseudo absence cell was placed randomly in the second order square of its respective presence cell (Figure 3). In cases where the two criteria were not met, the pseudo absence cell was allocated in the third or fourth order square. 80% of the pseudo absence cells could be placed in the second order ring.

This resulted in a pairwise selection of presence and corresponding pseudo absence cells and allows for a comparison of used versus unused habitat [45]. However, pseudo absence in a particular cell neither implies that the habitat is avoided, nor does this approach fully account for effects of population density and demography on habitat use. If, during the habitat survey, there was evidence of hazel grouse presence in form of sightings and/or indirect evidences (e.g. droppings, feathers), the species data was adjusted such that an affected pseudo absence cell was changed to be a presence cell. The resulting sample size of the field survey is summarised in Table 1.

Table 1: Field survey sample size per study region. The number in parenthesis indicate the adjusted presence cells due to evidences in the field survey.

Region	Presence cells	Pseudo absence cells	Total
Jura	35 (6)	27	62
Northern Prealps	31 (4)	26	57
Central Eastern Alps	28 (6)	19	47
Total for study area	*94 (16)*	*72*	*166*

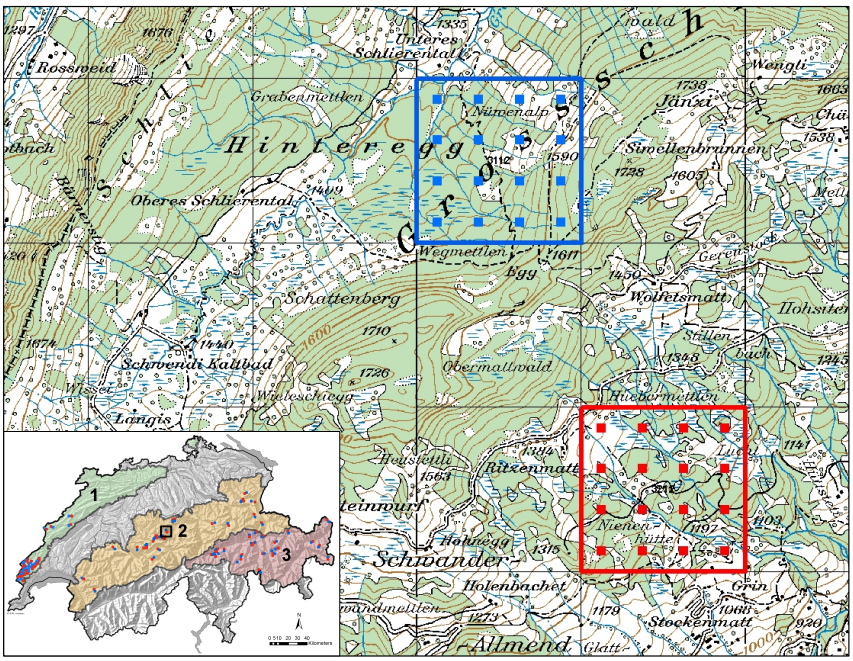


Figure 3: Sampling design indicating the km^2 grid of the Swiss coordinate system (black lines). A presence (blue) and a pseudo absence (red) sampling cell including 16 sampling plots is shown.

2.3 Habitat assessment

2.3.1 Field survey

In order to make collection of habitat data logistically feasible, four teams including a Master-student and field assistants were built to collect habitat data for all four target species in one study region (see 1.1 on page 2) . Once the data for all regions was collected, it was exchanged such that each student got the data for the assigned target species (i.e. hazel grouse in this case) from all regions. Field surveys were carried out between June and October 2011, while grouse presence grids were not sampled before mid July, thus minimising negative effects on juvenile development.

Each sampling cell was represented by 16 regularly distributed sampling plots (plots) of 30 x 30m (Figure 3 and 4). Habitat surveys were only carried out on plots that were covered with forest. The forest/non-forest decision was based on the decision scheme provided by the Swiss National Forest Inventory (NFI 3, [48]). It involves a set of decision criteria out of which the degree of cover (i.e. >20%), the spatial extent (i.e. width of interpreted area > 25m) and the minimal stand height (i.e. >3m) comprehend the main prerequisites for a forest decision. In cases where the plot area included an inner or outer forest edge (see below for definition) and its centre point was situated in a non-forest area, the plot was shifted into the forest so that the whole plot area was covered by forest. When a plot included an inner or outer forest edge and its centre was located in forest area, only the area covered by forest was considered for the habitat survey. Plot centres were located with a GPS device and marked with a ranging pole. The plot outline (30 x 30m) as well as the nested square were measured with a measurement tape and the corners marked with ranging poles. All plots were aligned parallel to the contour line with the top left ground vegetation plot being number 1 (Figure 4).

21 habitat variables were recorded considering hazel grouse' requirements in terms of forest composition and structure. According to species experts and literature (e.g. [27, 6, 92]) a combined set of habitat variables was developed to reflect: stand characteristics (4 variables), stand composition (2 variables), ground vegetation (10 variables), resources/functional trees (4 variables) and ecotones (1 variables).

The sampling reference within the plot area varies between the variables and is indicated in Table 2. Variables related to stand characteristics quantify the horizontal and vertical structure of the forest and thus reflect important habitat characteristics [27, 6]. Canopy and shrub cover are indicators for thermal and light conditions in the lower forest layers, therefore, they are expected to serve as a proxy not only for structural elements (e.g. cover from predators) but also for the availability of light demanding food plants [91, 6, 52]. Stand composition variables were included because they provide information about

Table 2: Field variables used to quantify habitat characteristics in a sampling plot (30m x 30m). For clarification regarding the sampling reference refer to Figure 4.

Variable	Definition	Operationalisation/Unit	Sampling reference
Stand characteristics			
Successional stage	Classification of the dominant stand's successional stage based on tree the height (h) and diameter at breast height (dbh, at 1.3 m above ground) distributions [48].	1 = regeneration (h $\leq$ 1.3m), 2 = thicket (h $\geq$ 1.3m, dbh $\leq$ 10cm), 3 = pole stage (dbh 11-30cm), 4 = tree-stage (dbh 31-60cm), 5 = old forest (min. 3 trees with dhb $\geq$ 61cm , 6 = multi-age forest	Plot (30 x 30m)
Stand structure	Classification of the vertical stand structure based on the proportion of different layers [48], multi-layered includes 'stufig' and 'Rottenstruktur'.	1 = one-layered 2 = two-layered 3 = multi-layered/tree-layered	Plot (30 x 30m)
Canopy cover	Area covered by crowns of trees (h $\geq$ 5m)	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Each quarter (15 x 15m) separately, averaged
Shrub cover	Area covered by shrub layer (h = 1.3 - 5m)	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Each quarter (15 x 15m) separately, averaged
Stand composition			
Canopy composition	Proportion of each species within the tree layer (h $\geq$ 5m) based on stem counts: Spruce (*Picea abies*), Fir (*Abies Alba*), Pine (*Pinus sylvestris & Pinus cembra*), Larch (*Larix decidua*), other conifers, Beech (*Fagus sylvatica*), resource trees (*Sorbus sp., Salix sp., Betula sp., Alnus sp., Corylus sp., Sambucus sp.*), other deciduous trees	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Plot (30 x 30m)
Shrub composition	Proportion of each species within the shrub layer (1.3m $\leq$ h $\leq$ 5m) based on stem counts of species listed in Canopy composition.	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Plot (30 x 30m)
Ground vegetation			
Ground vegetation cover	Area covered by ground vegetation (h $\leq$ 1.3m); includes all ferns, grasses, herbs, *Rubus sp., Vaccinium sp. Rhododendron sp., Juniperus sp. [48]*	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Ground vegetation plots 1 and 3 (7.5 x 7.5m), averaged
Ground vegetation distribution	Classification of the spatial distribution of ground vegetation	1 = clumped (vegetation restricted to one side of the plot), 2 = patchy (vegetation restricted to two large patches), 3 = evenly (several small patches, evenly distributed)	Plot (30 x 30m)
Cover of coniferous regeneration	Relative proportion covered by coniferous regeneration	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Ground vegetation plots 1 and 3 (7.5 x 7.5m), averaged
Cover of deciduous regeneration	Relative proportion covered of deciduous regeneration	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Ground vegetation plots 1 and 3 (7.5 x 7.5m), averaged
Cover of herbs	Relative proportion covered with herbs	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Ground vegetation plots 1 and 3 (7.5 x 7.5m), averaged
Cover of ferns	Relative proportion covered with ferns	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Ground vegetation plots 1 and 3 (7.5 x 7.5m), averaged
Cover of grass	Relative proportion covered with grasses	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Ground vegetation plots 1 and 3 (7.5 x 7.5m), averaged
Cover of berries	Relative proportion covered by berries excl. bilberry (*Vaccinium myrtillus*)	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Ground vegetation plots 1 and 3 (7.5 x 7.5m), averaged
Cover of bilberry	Relative proportion covered by bilberry (*Vaccinium myrtillus*)	Estimated in %-steps: 0, 1, 5, 10, ..., 95, 100	Ground vegetation plots 1 and 3 (7.5 x 7.5m), averaged
Ground vegetation height	Height of dominant ground vegetation	Measured in cm	Ground vegetation plots 1 and 3 (7.5 x 7.5m), averaged
Resources/Functional trees			
Total Rowans	Number of Rowans (*Sorbus aucuparia*), h $\geq$ 3m	Counted number	Nested square (15 x 15m)
Rowans fructifying	Number of fructifying Rowans (*Sorbus aucuparia*), h $\geq$ 3m	Counted number	Nested square (15 x 15m)
Basal-branched trees	Number of basal branched trees or groups of trees, i.e. trees (h $\geq$ 5m) with branches reaching down to 50cm or lower above ground	Counted number	Plot (30 x 30m)
Ant hills	Total number of ant hills	Counted number	Nested square (15 x 15m)
Ecotones			
Ecotones	Presence/absence and number of ecotones	E1=inner forest edge, E2=outer forest edge, E3=between successional stage, E4=along waterbodies, E5=along forest roads, E6=other	Plot (30 x 30m)

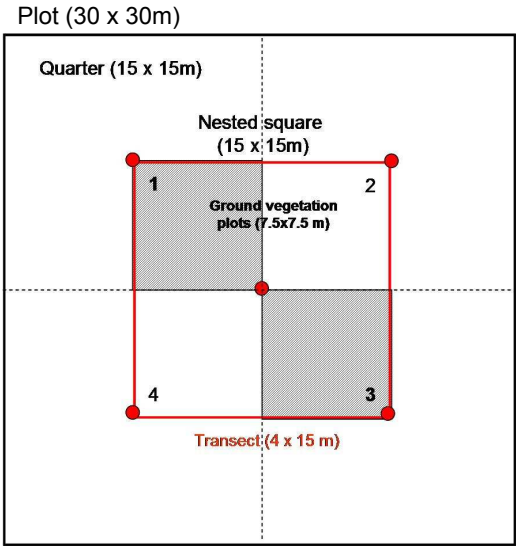


Figure 4: Plot sampling design [63]. The sampling plot (30 x 30m) is divided into 4 quarters (15 x 15m) of which the centre points define the corners of the nested square (15 x 15m). The nested square is again divided into 4 quarters (7.5 x 7.5m) of which the number 1 and 3 was used to sample the ground vegetation. The red transect (4 x 15m) was used to quantify lying deadwood, a variable not considered here.

the availability of resource trees and resource shrubs. Ground vegetation aspects were sampled in more detail because hazel grouse is a ground-nesting bird that depends on the availability of small-scale changing habitat elements providing visual cover and food resources [91]. Tall and fructifying rowans (*Sorbus aucuparia)* are an important resource related habitat requirement in Swiss mountain forests [15, 79]. The number of fructifying rowans taller than 3m as well as the total number of rowans above 3m were recorded. In this way, it was anticipated to sample a crucial variable - tall, fructifying rowans - while making it possible to overcome potential variation derived from regional and seasonal differences in the fructification intensity of this plant species. Ecotones were classified into 6 categories (Table 2). Inner forest edges were defined as transitions to gaps or clearings with a minimum area of 30 x 30m. Outer forest edges include transitions to open land. Further, ecotones between successional stages, along waterbodies and along forest roads were recorded. Other ecotones, the last category, include for example transitions to erosion gullies or steep hillsides.

2.3.2 LiDAR based habitat assessment

2.3.2.1 Technology Airborne laser scanning (ALS) or light detection and ranging (LiDAR) is an active remote sensing technology that directly measures the three-dimensional coordinates of an object in a reference system. A LiDAR sensor emits short

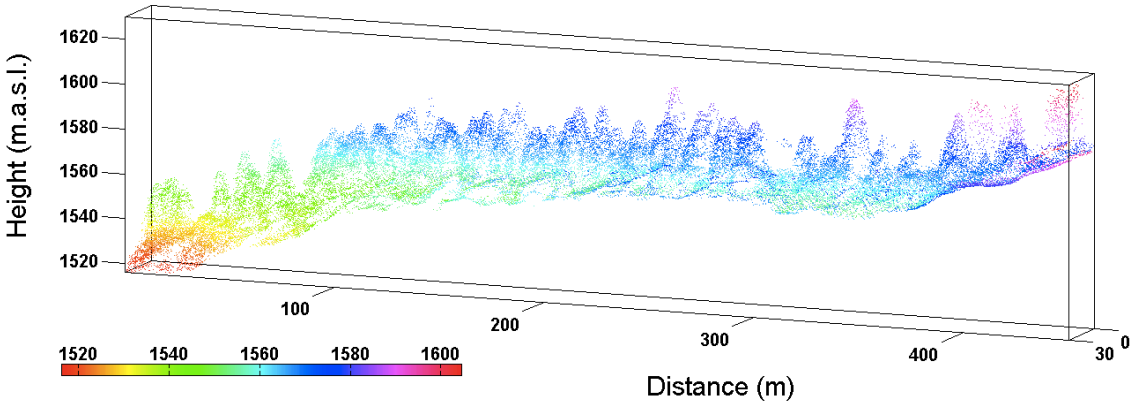


Figure 5: Three-dimensional representation of a data point cloud consisting of DOM-AV and DTM-AV points.

Table 3: Key characteristics of Swisstopo LiDAR data.

Average flight height	1200 m
Flight season	all year round
Areal coverage	Switzerland; all locations below 2000 m.a.s.l.
Average point density	0.5 point/m^2
Nominal footprint size	in the range of small-footprint laser scanning (several decimetres)
Standard error in height accuracy	50cm in open areas 150cm in forest areas
Source	DTM-AV, DOM-AV © 2011 Swisstopo (5704 000 000)

laser pulses to the earth's surface and measures the elapsed time from the emission to the detection of the reflected pulse. This time span multiplied by the speed of light allows for a precise calculation of the distance to the reflecting object. The exact position of a reflection is then derived by combining measurements of the sensor's position and alignment. The direct product of these measurements is a point cloud consisting of points (i.e. laser echos) with x,y and z-coordinates [3, 88]. Figure 5 illustrates an example of such a point cloud. It can be seen that beside points reflected by the vegetation, some points were also reflected by the terrain, which is an important discrimination for deriving further products from the point cloud.

Table 4: Sample size of LiDAR based habitat assessment per study region.

Region	Presence cells	Pseudo absence cells	Total
Jura	35	27	62
Northern Prealps	26	23	49
Central Eastern Alps	30	20	50
Total for study area	*91*	*70*	*161*

2.3.2.2 Swisstopo data Between 2000 and 2007 the Federal Department of Topography (Swisstopo) acquired first and last return LiDAR data to form a nationwide dataset. During this period, several companies were involved in the data acquisition and subsequent preprocessing, which resulted in a variation in documentation and properties (e.g. point density) within the dataset. Table 3 summarises some key characteristics of the data.

The data provided by Swisstopo is divided into two parts: DOM-AV and DTM-AV. This is a result of filtering and classification algorithms that were applied in the preprocessing of the data. The DOM (Digital Surface Model) file consists of points classified as vegetation, artificial structures or buildings, and terrain. The DTM (Digital Terrain Model) file contains only points classified as terrain and is able to provide a more detailed representation of the terrain in forested areas compared to the DOM-file (R. Artuso, Swisstopo, personal communication). Classification accuracy lies between 80 and 95% as reported by the data provider. However, since this study concentrates on mountain forests with relatively few artificial structures and buildings, the classification accuracy of the used data in the DOM-file is expected to be higher. With respect to the point density for the areas analysed here, the range of density for all points in the DOM-file is 0.74 to 2.80 points/m^2 with a mean of 1.38 points/m^2.

For the perimeter of each sampling cell, all DOM and DTM points were extracted at the Swiss Federal Research Institute WSL. In order to allow for edge detection at the border of the applied perimeter, a buffer of 50 meters was added. This resulted in the data extraction for cells measuring 1100 x 1100 meters. Unfortunately, the DOM-AV has not been classified for an extended area in the central part of Switzerland, which affected the study region Northern Prealps. As a result, 11 sampling cells in the Northern Prealps could not be processed, however, this loss was partially compensated by the availability of LiDAR-data for 6 sampling cells that were not accessible and thus not sampled within the field survey. The number of cells processed is indicated in Table 4.

2.3.2.3 Data processing The LiDAR data was processed in MATLAB (Version R2011a). This software is particularly strong at performing computationally intensive tasks which is a prerequisite for this study, considering the large amount of data at hand.

The idea was to create a routine that executes all steps for each sampling cell auto-

matically, from the import of data, the processing of the data to the export in the desired formats. This has the advantage that processing becomes efficient and can easily be repeated for similar studies. Many of the parameters were extracted using routines originally written by Dr. Felix Morsdorf (personal communication) and developed in the course of several research projects (e.g. [68, 54, 53, 69]). Integrated MATLAB functions complemented the processing. Figure 6 shows a flow-chart indicating all of the major processing steps. It is accompanied by the applied and commented MATLAB code in Appendix B. The numbered steps in the flow-chart are explored in detail in the following explanation.

After loading the DOM-AV and DTM-AV-file (1), all points classified as vegetation were extracted from the DOM-AV (2). All other DOM-point classes were excluded from further analysis. In order to obtain a continuous surface of the terrain, the DTM-points were processed to form a digital terrain model. The function used applies a moving window algorithm to determine each raster cell value. The final terrain raster was calculated with a resolution of 1m, considering all points available within a moving window perimeter of $4m^2$ (3). In areas with dense vegetation, the point density in the DTM-AV-file can be lower than average, which results in small gaps of the terrain raster. These gaps were closed using a linear interpolation algorithm. Having filtered for vegetation points and calculated a digital terrain model, the terrain corrected height and position of each vegetation point was determined by subtracting the terrain height from the vegetation height (4). In Figure 7, all terrain corrected vegetation heights are shown for a sampling cell of $1km^2$, together with a transact indicating the vertical profile. The terrain corrected vegetation point cloud allows for an important product of LiDAR data: the Canopy Height Model (CHM). It was computed using the moving window algorithm mentioned above, setting the resolution at 4m with a window perimeter of $64m^2$. The decision for a 4m resolution was based on the assumption that $16m^2$ is an acceptable approximation of a single tree crown projection in spruce dominated mountain forests [44]. In order to allow for further analysis in FRAGSTATS, the CHM was classified into the following four classes (6):

1. Matrix, no vegetation points
2. Understory class, vegetation heights of 5m and lower
3. Midstory class, vegetation heights between 5m and 15m
4. Canopy class, vegetation heights of 15m and above

Figure 8 contains two examples of a CHM representing a presence and a pseudo absence cell. The rational behind the height classification was to define an ecologically meaningful stratification of the forest while keeping it as simple as possible. The classified CHM was

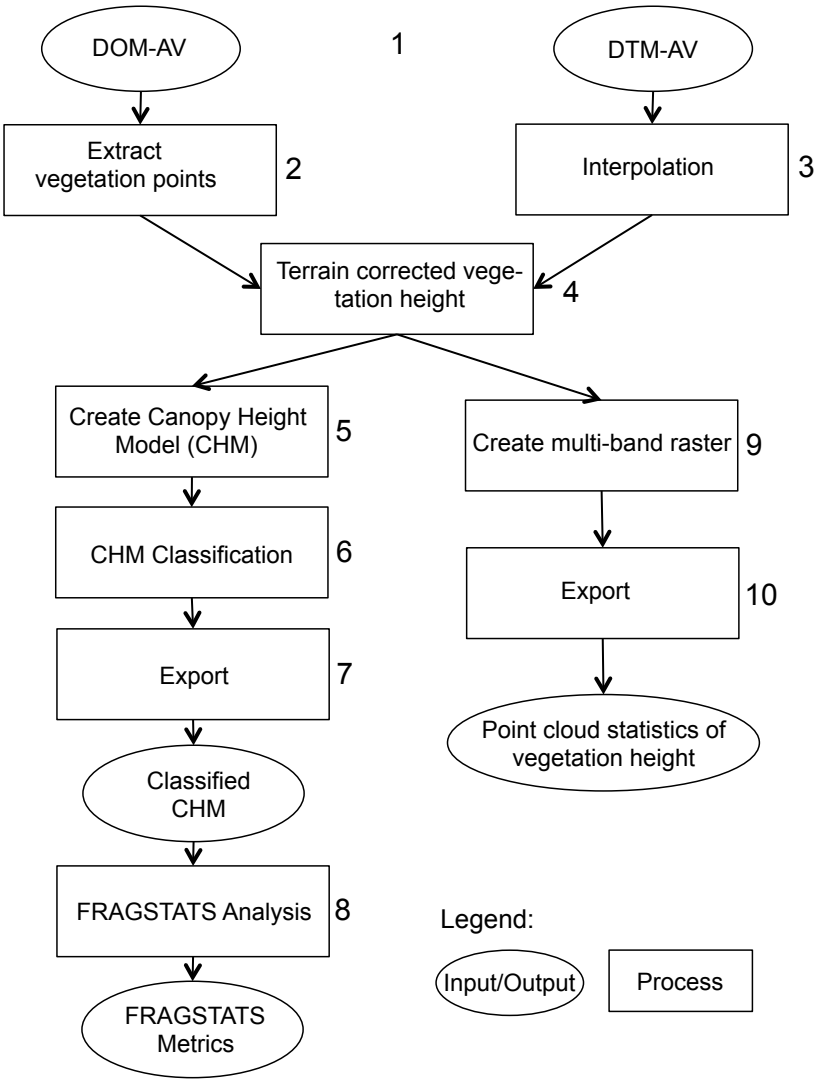


Figure 6: Flow-chart for processing LiDAR data. DOM-AV: raw-data point cloud, classified as vegetation, artificial structures or terrain; DTM-AV: raw-data point cloud, classified as terrain (2.3.2.2). The terrain corrected vegetation height is obtained by subtracting the interpolated terrain heights (3) from the raw-data heights of the vegetation points (2). Further details for each numbered processing step are provided in the text (2.3.2.3).

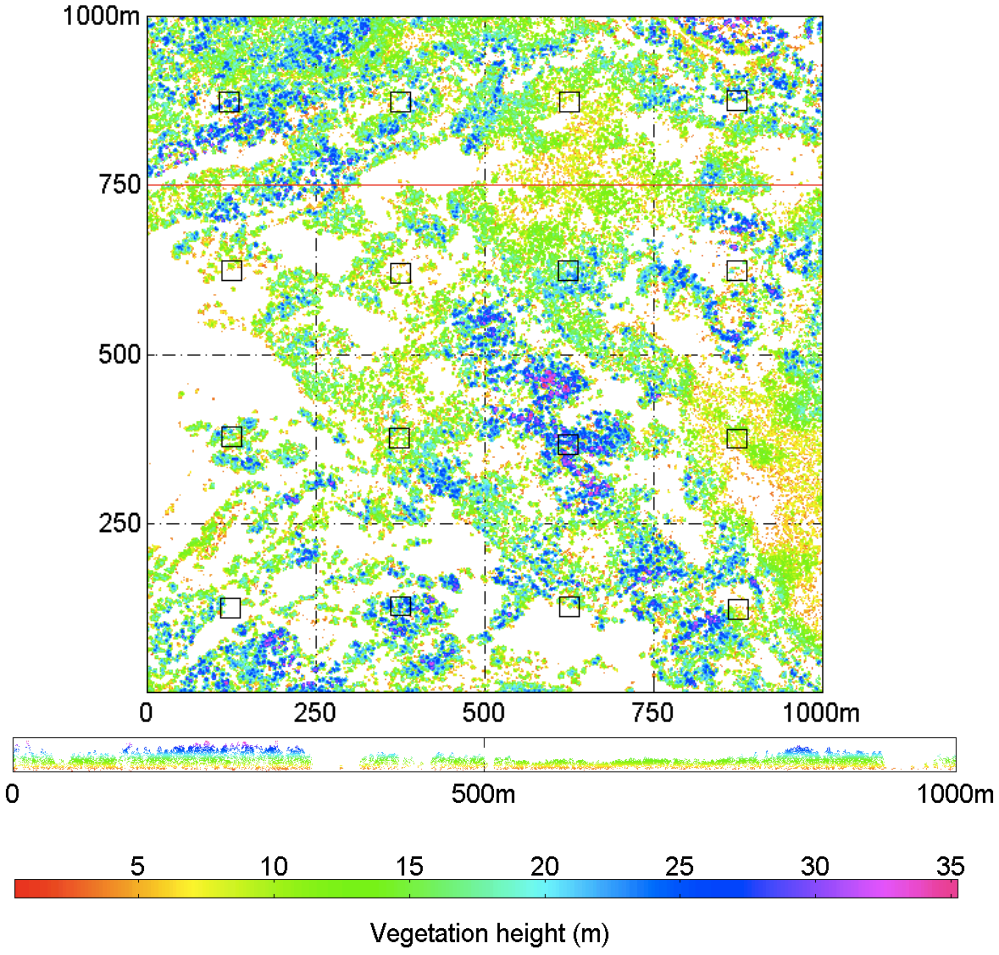


Figure 7: Orthogonal projection of terrain corrected vegetation heights for a sample cell of 1km^2. An example of a profile is illustrated along a transect (red line at y=750). The nested squares indicate the 30x30m field sampling plots.

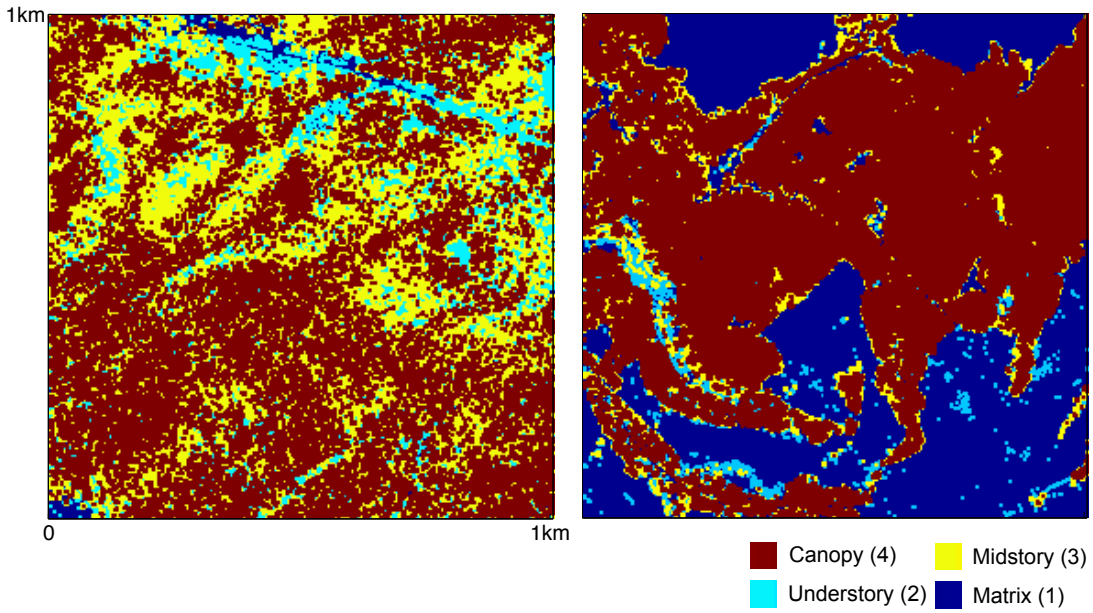


Figure 8: Classified Canopy Height Models (CHM) showing a presence (left) and a pseudo absence (right) km^2 cell. Raster-values (4 x 4m cell size) are based on interpolated vegetation heights. Canopy class (4) = vegetation heights $\geq$ 15m; Midstory class (3) = vegetation heights $>$ 5m and $\leq$ 15m; Understory class (2) = vegetation heights $\leq$ 5m; Matrix (1) = no vegetation heights.

exported as an ASCII-grid (7). Further analysis of the classified CHM in FRAGSTATS (8) are explained in paragraph 2.3.2.5.

Apart from calculating the CHM, the terrain corrected vegetation heights were also used to derive a multi-band raster of the forest (9). The multi-band raster was created with a horizontal resolution of 10m, where the vegetation points were subdivided into 10 bands or layers. Applied to the sampling cell of 1km^2, a horizontal resolution of 10m results in a raster of 100 x 100 cells with the vegetation points being subdivided into 10 vertical bands for each cell (Figure 9). For every sampled km^2, the means of the statistical components min, max, mean and standard deviation per band were calculated based on the band values per cell. The band values are computed based on two methods: height and density. Height related band thresholds show a varying distribution across the height profile. The first (lowest) height related band consists of the lowest 10 percent of the vegetation points (i.e. the 10th percentile). Therefore, the upper border of the first height band in an open forest with a well developed understory (many vegetation points in the understory) will be lower than the upper border of the first height band in a forest with a closed canopy and hardly any understory (few points in the understory). On the other hand, density related band thresholds are defined by a fixed height distribution. For example (see Figure 9), a forest with the tallest tree being 37m high will be divided into the defined number of height

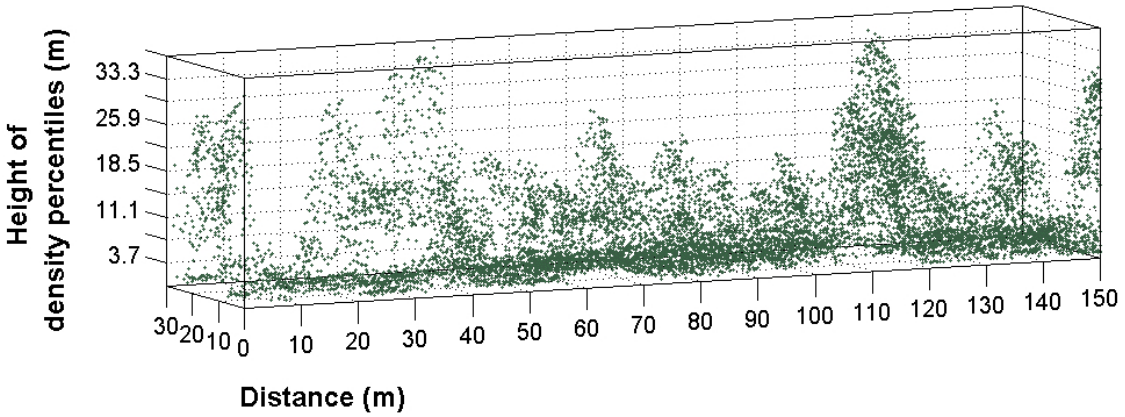


Figure 9: An extract of a multi-band raster of terrain corrected vegetation point heights. The x- and y-axis form a raster grid with a resolution of 10m, whereby the z-axis is split into 10 bands or layers. In this case, the bands are the density percentiles with a fixed height distribution from the 10th percentile (3,7m) to the 100th percentile (37m) (see 2.3.2.3 for details).

bands (e.g. 10) which results in height thresholds of 3.7, 7.4, 11.1, ..., 37m. The point density in each of these fixed band heights will then be used to calculate the density related statistics. The underlying function had to be adapted to fit the data[2]. A cell size of 10m was considered reasonable because it provides for a sufficient amount of points available for the calculations, i.e. more than 100 points per cell. 10 layers were chosen to indicate the percentiles, and thus to render a detailed statistical picture of the vertical vegetation point distribution within the forest. It was therefore expected to give a reasonable compromise between horizontal resolution and the informative value in the vertical dimension. The output of the multi-band raster was exported as comma-separated values (10). Together, these point cloud statistics and the classified CHM served as a basis for deriving variables related to habitat characteristics, which is the content of the next paragraphs.

2.3.2.4 Statistical variables from point cloud Table 5 summarises all variables derived from the terrain corrected vegetation point cloud. The penetration rates were extracted to serve as an indicator of light and thermal conditions in the shrub layer. The standard deviation of the height and density percentiles provide information about the variability within a particular band. The aim was to find a proxy for structural heterogeneity, which is expected to be an important habitat element [27]. The mean values of the height and density percentiles contain information about the stratification of the vegetation height. Crucially, the mean of the lower percentiles were assumed to reflect how well the

[2]Since the data contained no information about the intensity of the laser echos, all intensity related elements of the code had to eliminated.

Table 5: Predictor variables derived from the vegetation point cloud.

Variable	Definition
Penrate1-4m	Penetration rates on four height levels: 1m, 2m, 3m, 4m. Expressed as: $Sum\,of\,points \leq height\,level / sum\,of\,all\,points$
$StdH_{10,\ldots,100}$	Standard deviation for each height percentile
$StdD_{10,\ldots,100}$	Standard deviation for each density percentile
$MeanH_{10,\ldots,100}$	Mean for each height percentile
$MeanD_{10,\ldots,100}$	Mean for each density percentile

understory layer is developed. Due to the lack of knowledge and experience regarding the ability of these variables to describe important habitat characteristics, all ten percentiles were derived in order to allow for an exploratory approach considering all data available. Finally, all variables were spatially averaged to a single value per sampling cell (1km^2).

2.3.2.5 FRAGSTATS metrics FRAGSTATS is an open source software designed for the analysis of spatial patterns and quantification of landscape structure [66]. Although initially developed for landscape ecologists, FRAGSTATS can be used to analyse spatial phenomena in any categorical map. In this thesis, the tool was used to quantify spatial patterns in the classified CHM (2.3.2.3). Data import was performed using a batch-file, which allows for an efficient handling of multiple landscapes (i.e. the CHM-ASCII-grids of all sampling cells). In FRAGSTATS, literally hundreds of indices of landscape patterns can be computed, however, the metrics considered most relevant for this study are summarised in Table 6. The selection of relevant metrics was based on the degree of contrast obtained by a test phase in which a wide set of metrics were tested on a subsample of the data. The aim was to cover several ecological aspects such as the density of the forest vegetation, its distribution in the landscape and the quantification of ecotones within the landscape. It is important to note that FRAGSTATS computes metrics on different levels, i.e. patch, class and landscape. Patch-level metrics describe the spatial character and context of individual patches within a landscape. However, the patch metrics in this analysis were primarily used to compute several class- and landscape-level metrics, for example by averaging the patch attributes across all patches of a given class. Class-level metrics quantify the amount and spatial configuration of all patches of a given type or class. Therefore, they serve as a means to quantify the extent and fragmentation of a patch class in the landscape. Landscape-level metrics are integrated over all classes over the full extent of the landscape and thus allow for the analysis of the composition and configuration of entire landscape mosaics [66]. In

Table 6: Predictor variables based on FRAGSTATS metrics. CHM classes: matrix (1), understory (2), midstory (3), canopy (4).

Variable/Metric	Level	Unit	Definition
ShDI2_3	landscape	index	Shannon's diversity index for CHM classes 2 & 3
ShDI1_2	landscape	index	Shannon's diversity index for CHM classes 1 & 2
ShEI2_3	landscape	index	Shannon's evenness index for CHM classes 2 & 3
Area_Am_2-4	class	area (ha)	Area-weighted mean of patch area for CHM classes 2, 3 and 4
Area_Cv_2-4	class	area (ha)	Coefficient of variation of patch area for CHM classes 2, 3 and 4
Frac_Mn_2-4	class	index	Mean fractional dimension index of CHM classes 2, 3 and 4
Clumpy_2-4	class	index	Clumpyness of CHM classes 2, 3 and 4
Te1_2	landscape	length (m)	Total edges of CHM classes 1 and 2
CHH	landscape	length (m)	Total edges of CHM classes 2, 3 and 4; Canopy Height Heterogeneity (CHH)

this study, only metrics at the class and landscape level were considered (Table 6).

Shannon's diversity index was calculated for the understory and midstory as well as the matrix and understory class because diversification between these classes was expected to be most relevant for hazel grouse. This index is influenced by two components - richness and evenness, and thus refers to compositional and structural aspects of the CHM [66]. However, Shannon's diversity index is quite sensitive to richness, which is, in this case, the number of different classes (i.e. two). Therefore, the Shannon's evenness index for the understory and midstory class was derived separately. The area-weighted mean patch area for each class was included due to inherent information about the spatial extent and distribution of the forest vegetation. To reflect variability among the patch areas, the coefficient of variation was derived for each class. A fractional dimension index quantifies the degree of complexity of the shape of a particular patch. More complex patch shape implies more transitions to patches of other classes and could thus be relevant for quantifying ecotones in a landscape. Clumpyness describes the extent of aggregation of a CHM class [66] and was computed for each class.

The total amount of edges for the matrix and understory class is expressed by the sum of the lengths (m) of all edge segments involving the two patch classes. Since a resolution of 4m was defined to derive the CHM (2.3.2.3), each adjacent pixel from another class increases the amount of edges by 4m. The same principle applies to the total edges of the understory, midstory and canopy class. The aim here was to quantify heterogeneity within the forest vegetation, the matrix was therefore excluded from the calculation of this metric. Thus, the total length of edges of the understory, midstory and canopy class are referred to as Canopy Height Heterogeneity (CHH). The boundary segments of the landscape, i.e.

the border of the 1km^2 sampling cell, were not included in any edge calculation. Further, all metrics in Table 6 are regarded as predictor variables and as such included in the model fitting and evaluation.

2.4 Statistical/Machine learning analysis

2.4.1 Boosted regression trees

Boosted regression trees (BRT) were used to model the relationship of hazel grouse presence and pseudo absence with respect to field and LiDAR derived predictor variables. BRT is a machine learning method which has only recently started to be applied in ecology [35, 24, 19, 57]. The name of this modelling technique is based on the integration of two algorithms: regression trees and boosting. Regression trees (or decision trees) apply rules to identify regions with most homogeneous responses to predictors, and fit the mean response for observations in each region [19]. Regression trees are an advantageous technique for ecological modelling because predictor variables can be of any type (numeric, binary, categorical, etc.), do not need to be transformed, and irrelevant predictors are largely ignored. Furthermore, regression trees are insensitive to outliers, deal with missing data in predictor variables and model interactions between predictors automatically [19]. The downside of regression trees lies in the fact that they are often not very accurate, which is where boosting comes into play. Boosting is a forward, stage-wise procedure that improves model accuracy by progressively building a sequence of models (i.e. regression trees). Each subsequent tree concentrates on explaining the residuals of the previous trees, thus slightly decreasing the deviance of the full model [19]. Improving the accuracy of a boosted model can further be achieved by introducing some randomness or stochasticity, however, this introduces variance in fitted values between runs [24, 19]. In BRT, the level of stochasticity is controlled by the 'bag fraction', which defines the size of a random subset of the data on which each tree is fitted [19].

There are two inter-dependent features of BRT: the learning rate (lr) and the tree complexity (tc). The learning rate or shrinkage controls the contribution of each tree when it is added to the model. A small value (e.g. 0.0005) implies that the contribution of each tree is shrunk less than with a high value (e.g. 0.5). The tree complexity (sometimes referred to as interaction depth [57]) defines the number of nodes in a tree. A tc value of two or more implies that variable interactions effects are fitted [19]. Together, lr and tc determine a third parameter for optimal model performance, the number of trees (nt). To control the BRT model for over-fitting, a regularisation method that jointly optimises the values for lr, tc and nt is applied. This method aims at minimising the deviance and uses cross-validation (CV) to identify optimal settings. The procedure is explained in detail by [19]. Beside for model development, CV is also used in model evaluation. The CV

method divides the whole dataset into 10 subsets out of which 10 different training sets consisting of a unique combination of 9 subsets are created. Each training set can then be tested on the omitted unique subset. The fitted values based on this 10-fold CV procedure provide a solid base for evaluating the predictive performance or level of generalisation of a particular BRT model [19, 33].

Apart from the ability of BRT to model typical ecological features such as nonlinearities and interactions, comparison studies also revealed that BRT outperform more traditional modelling techniques such as GLMs and GAMs not only in relation to explaining the observed data but even more so in terms of predictive performance [17, 16, 57, 58]. These findings as well as the availability of properly documented code made it a promising approach to fit and evaluate the models with the data at hand.

2.4.2 Model fitting and evaluation

All BRT related analysing steps were carried out with R [1] version 2.13.2 using the dismo package [36] and additional custom written code provided by [19] and [18].

Although BRT effectively selects relevant predictor variables, it is recommended by [19] to simplify the set of predictor variables, especially in small data sets. In a first step, bivariate correlations were explored for the field and LiDAR derived variable sets separately, since they were regarded as independent from each other. A correlation coefficient higher than 0.7 was applied as a threshold value for variable exclusion [21]. In cases where the Spearman rank correlation coefficient of two variables exceeded 0.7, the variable considered to be ecologically more meaningful was retained. In a second step, univariate generalised linear models were calculated for all predictor variables using the GLM function in R. Only significant variables ($p \leq 0.05$) were included in further analysis. A descriptive summary of all significant variables is provided in Appendix A. Although this second step is not common in fitting BRT models, it provided an informative overview of the individual explanation potential of each variable. Furthermore, BRT models were run with and without this second variable reduction step and the results revealed hardly any differences in terms of retained variables, with a tendency towards the more simplistic variable sets performing better. The third step was based on the simplifying procedure as proposed in the supplement material provided by [19] and involves a sequential removal of least important predictor variables while calculating the change of predictive deviance for each sequence (i.e. for each new set of variables). As can be seen in Figure 10 (which is based on the combined model listed in 3.1), dropping of up to 6 variables would hardly have any effect on the predictive performance. After 7 dropped variables, the predictive deviance decreases, which would result in a better performing model. The lowest change in predictive deviance is reached at 10 dropped variables, after that, the model performance decreases again. Hence, the 10

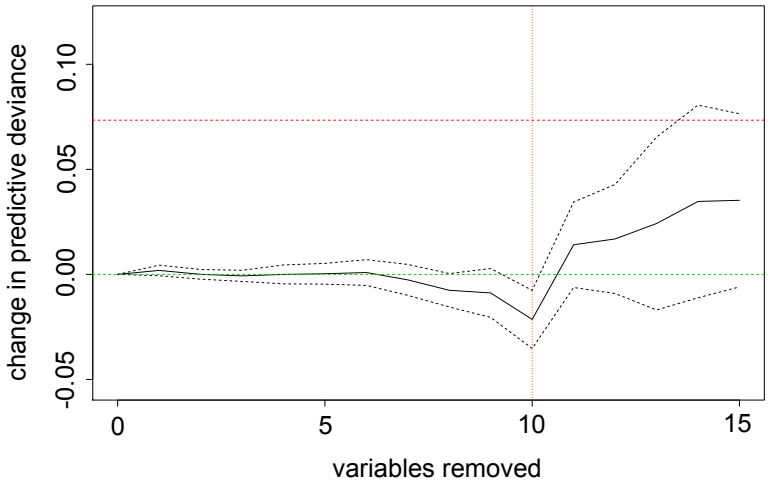


Figure 10: Simplification of a BRT model indicating that the model performance could be improved by dropping 10 variables. The solid line shows the mean change in predictive deviance, and the dotted line one standard error based on 10-fold CV (Appendix 2 in [19]).

least important variables were dropped and the whole model was recalculated using the reduced set of variables. This procedure was applied to all fitted final models.

Three final models were fitted: a field model considering the field variables, a LiDAR model considering the LiDAR variables and a combined model with field and LiDAR data taken together. To identify the optimal settings for *lr, tc* and *nt* each model was run using a combination of different settings. These parameter settings included a range of *lr* values from 0.5 (fast) to 0.0001 (very slow) in combination with values for *tc* of 1 (no interaction), 2, 3 and 4 (4-way interaction). Bag fractions of 0.5, 0.75 and 0.9 were tested on a subsample of the models and the outcome suggested a bag fraction of 0.75, which is in line with recommendations by [19] and [57]. The same authors state that, as a rule of thumb, the values for *lr* and *tc* should allow for an optimal number of trees of 1000 and more, which was taken into account when the final values for *lr* and *tc* were selected. As described above (2.4.1), applying a bag fraction of 0.75 improves model accuracy and introduces a stochastic element to the model fitting. This implies slightly differing results for each run depending on the randomly selected subset of the data. This variance was accounted for by taking the means of 5 repetitions of each model [57].

The code provided by [19] also allows for the calculation of the relative importance of a predictor variable. It is computed based on how often a variable was selected, and how much the model improved as a result of the selections. The relative influence of all variables retained by the model adds up to 100% [13]. Furthermore, another function in the dismo package was used not only to identify but also to quantify important interactions between predictor variables [19]. To serve as a basis for interpretation of the results, the

effect of each predictor variable on the response variable is visualised in the form of partial dependence plots. These plots indicate the relative importance of a predictor variable after taking into account the average effect of all other variables in the model [35]. Partial dependence plots can therefore be used to identify thresholds within a particular variable.

Evaluation of model performance was based on training data and predictions carried out on test data using CV. Thus the ability of the models to explain the observed data as well as to predict left-out data can be analysed. Two evaluation measures were calculated: the percentage of explained deviance (D^2) and the area under the receiver operating characteristic curve (AUC). The percentage of deviance explained is calculated as the difference between the deviance of a null model and the explained deviance, divided by the null model deviance. It is the equivalent to R^2 in least square regression [33]. The AUC value indicates the model's ability to discriminate between presence and pseudo absence and allows for a performance comparison between models [20]. The interpretation scheme proposed by [40] states that an AUC value of 0.7-0.8 can be regarded as an acceptable model performance, 0.8-0.9 is excellent and higher than 0.9 is considered outstanding. A value of 0.5 indicates a discrimination ability attained by chance.

2.4.2.1 The variable altitude and regional aspects Within the above described simplification process, altitude was found to be a significant variable and was listed in Appendix A. However, altitude is an abiotic factor that cannot be amended by forest management practices, furthermore, the contrast in altitude is influenced by the sampling design. Hence, it is questionable whether this variable should at all be included in this habitat analysis. The results and discussion therefore concentrate on models that were fitted without altitude.

The univariate testing of variables in Appendix A revealed another aspect inherent in the data: the study regions Jura and Northern Prealps responded much more to the variables than the Central Eastern Alps. This suggests a regional analysis, which is, however, beyond the capacity of this thesis. Nevertheless, the quantitative influence of altitude and exclusion of the Central Eastern Alps on model performance was assessed. The results are listed in 3.5 but not taken up in the discussion.

3 Results and interpretation

Beside the models performance assessment, it is crucial to a habitat analysis to know the contribution of each predictor variable. Therefore, the variable contributions are listed separately for each model. The same applies to the threshold values, which are reported for the most important predictor variables. Furthermore, variable interactions were identified, fitted and listed, however, interactions were weak and thus not further analysed.

3.1 Model performance

The field and combined model performed best when the set of variables were reduced as described in (2.4.2), where the field model retained 6 and the combined model 11 variables (Table 7). The LiDAR model retained all 9 variables listed in Appendix A and lost some of its ability to explain the observed data when further variables were dropped. As can be seen in Table 7, all three models fitted more than 1000 trees. The slow learning rate of the combined model, together with the tree complexity of 4, resulted in the highest number of trees fitted. In terms of explaining the observed data (D^2), the field model rendered slightly better results than the LiDAR model. The area under the receiver operating characteristic curve (AUC) based on the training data was about the same in the two models. Regarding the prediction of left-out data based on 10-fold cross-validation, the field model also performed better than the LiDAR model, however, the AUC values based on cross validation did not differ much between the two models. The combined model, on the other hand, reveals the best results not only in terms of explaining the observed data and AUC values, but also in terms of predicting left-out data and AUC values based on cross validation.

Table 7: Optimal settings, performance and AUC values for best performing BRT models relating hazel grouse presence and pseudo absence to field and LiDAR based variables, as well as a combination thereof. Beside the model name and the number of variables retained by each model, the table indicates the number of trees fitted (nt), at which learning rate (lr) and tree complexity (tc), the deviance explained and AUC values based on the training data (D^2, AUC), the total deviance explained and AUC values based on 10-fold cross-validation (CV D^2, CV AUC). Indicated figures represent the mean of 5 repeated runs of each model. The bag fraction is 0.75 (2.4.2).

Model	**No. variables**	**nt**	**lr**	**tc**	**D^2**	**AUC**	**CV D^2**	**CV AUC**
Field	6	1750	0.003	1	37.02	0.892	21.09	0.796
LiDAR	9	1840	0.003	1	34.13	0.884	17.79	0.781
Combined	11	5370	$<$0.001	4	55.41	0.977	23.46	0.804

Table 8: Predictor variables contributions to the total deviance explained (based on 10-fold CV) for the field, LiDAR and combined model.

Variable name	**Contribution to model in %**		
	Field	LiDAR	Combined
Field variables			
Ground vegetation cover	20.07		10.64
Basal branched trees	19.60		7.76
Total rowans $\geq$ 3m	17.43		7.33
Ground vegetation distribution	16.11		6.59
Resource trees in shrub layer	15.57		11.54
Bilberry cover	11.21		8.20
Fern cover			7.42
Total	100		59.48
LiDAR variables - FRAGSTATS			
Canopy Height Heterogeneity (CHH)		44.20	19.55
Clumpy Index for CHM class 4 (Clumpy_4)		15.63	8.81
Shannon's Diversity Index for CHM class 1 & 2 (ShDI1_2)		3.13	5.70
Shannon's Diversity Index for CHM class 2 & 3 (ShDI2_3)		2.88	
Total		65.84	34.06
LiDAR variables - point cloud statistics			
Standard deviation in 40th density percentile ($StdD_{40}$)		12.36	
Standard deviation in 80th height percentile ($StdH_{80}$)		9.33	6.45
Standard deviation 10th height percentile ($StdH_{10}$)		8.24	
Mean 70th density percentile ($MeanD_{70}$)		3.13	
Mean 20th density percentile ($MeanD_{20}$)		1.10	
Total		34.16	6.45

3.2 Variable contribution

3.2.1 Field model

From all six variables, the ground vegetation cover explained most of the deviance (20%), however, a similar contribution comes from the total number of basal branched trees. The rest of the variables explain about the same amount of deviance, with the bilberry cover being the least important variable. Resource related variables (i.e. total rowans $\geq$ 3m, res.

trees in shrub layer, bilberry cover) accounted for about 44% of the deviance explained, structural variables, on the other hand, contribute about 56% to the deviance explained (Table 8).

3.2.2 LiDAR model

The FRAGSTATS variable canopy height heterogeneity contributed most to the deviance explained (44%), followed by the clumpyness of the canopy class (2.3.2.5) (16%). A further 30% came from the standard deviations within the 40th density and 80th and 10th height percentiles. Other variables derived from point cloud statistics include the means of the 70th and 20th density percentile, which together contributed a little bit more than 4%. The Shannon's Diversity Indexes between the CHM classes 1 & 2 and 2 & 3 contributed 6% to the deviance explained. Taking all FRAGSTATS variables together, they accounted for 66% of deviance explained, whereas the variables derived from the point cloud statistics (2.3.2.4) contributed the remaining 34% (Table 8). The total deviance explained is based solely on structural variables because LiDAR only measures structural vegetation aspects.

3.2.3 Combined model

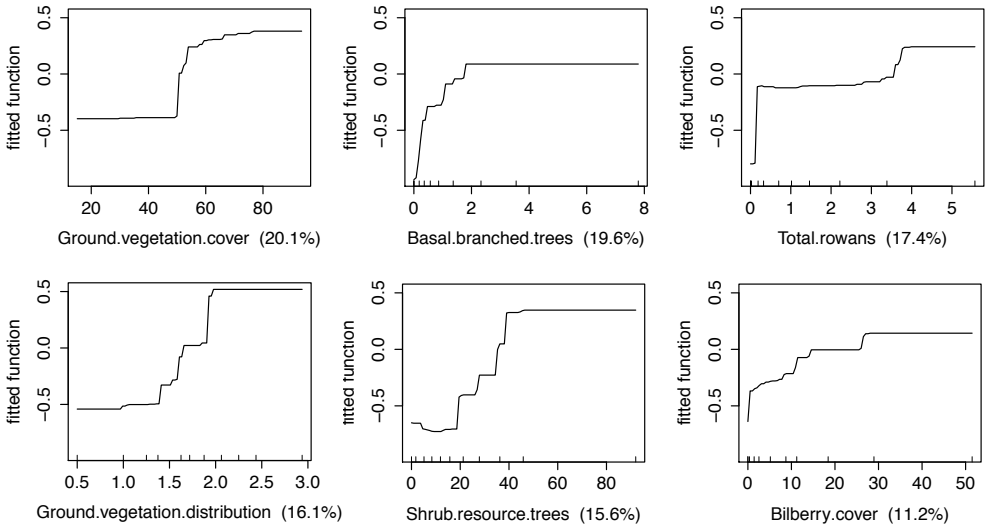
The canopy height heterogeneity contributed more than 19% to the explained deviance, which is 8% more than the cover of resource trees in the shrub layer and almost double compared to the ground vegetation cover. Thereafter, the contributions decrease more or less gradually (Table 8). Except for the LiDAR variables derived from point cloud statistics, the combined model retained most of the variables that were retained by the individual field and LiDAR models. All field variables together contributed almost 60% to the deviance explained, the remaining 40% comes from LiDAR variables. LiDAR variables based on FRAGSTATS perform much better (34%) than LiDAR variables derived from the point cloud statistics (6.45%). 73% of the contribution to the deviance explained derives from structural variables, 27% from resource related variables.

3.3 Threshold effects

3.3.1 Field model

The most obvious and most influential threshold for hazel grouse presence was a ground vegetation cover of 50% (Figure 11). Below 50% the response was negative, above it was positive and steadily increased with increasing ground vegetation cover. A minimum of two basal branched trees was required to achieve a positive response, however, it did not increase with more basal branched trees. The total number of rowans taller than 3m showed two thresholds, one just above zero, indicating that already a very low amount of

Field model



LiDAR model *(6 most influential variables)*

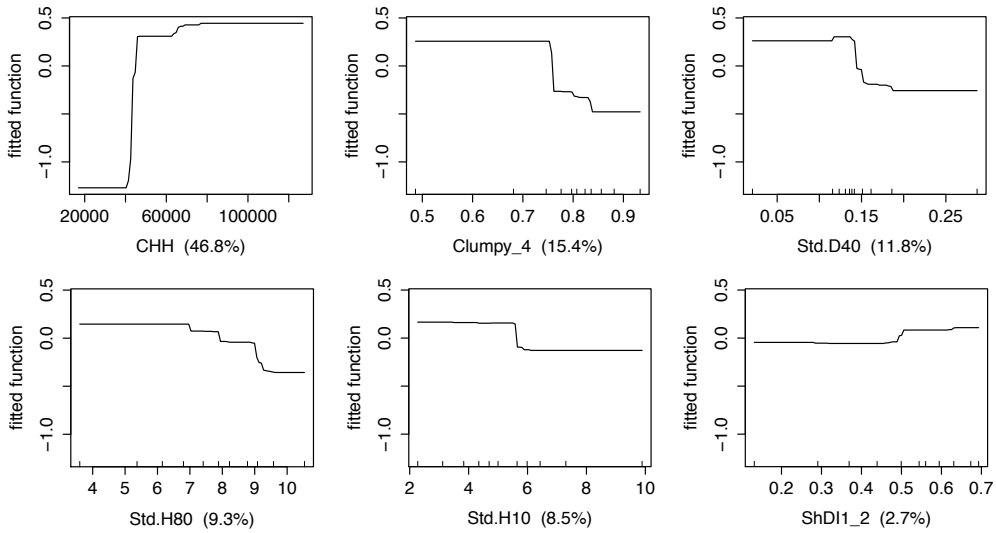


Figure 11: Partial dependence plots for variables retained by the final BRT field and LiDAR model. The graphs show the effect of a particular variable on the response (i.e. presence/pseudo absence): increasing fitted function values indicate that hazel grouse responded positively, decreasing values the opposite. Values on the y-axis are uniformly scaled so that the relative importance of the variables can be visually compared [35]. The relative contribution of each variable to the total deviance explained is shown in parentheses.

this resource tree positively affected the response, the other just below 4, which means that 4 or more tall rowans were able to improve the habitat quality for hazel grouse considerably. The graph for the ground vegetation distribution is difficult to interpret because it involves a categorical variable, however, a more evenly distributed ground vegetation (i.e. values closer to 3) had a positive effect on the response. The fitted function for the cover of resource trees in the shrub layer increased between 20 and 40%, and achieved a positive response at 35%. Between 0 and 28% cover of bilberry, the response increased, but did not further increase with values above 28%.

3.3.2 LiDAR model

The heterogeneity of the canopy height (CHH) was a strong indicator that had a positive effect on the response with a clear threshold value of 40'000 (Figure 11) . Low values for clumpyness of vegetation heights above 15m (2.3.2.3) rendered a positive response, however, values above 0.75 decreased the habitat quality. Thresholds within the standard deviations of the 40th density as well as the 80th and 10th height percentiles were indicated less strongly. But generally, higher standard deviations within these percentiles revealed a negative response. The Shannon's Diversity Index for the CHM classes 1 & 2 (2.3.2.5) was a rather weak indicator and affected the response slightly positive at and above a value of 0.5.

3.3.3 Combined model

The three most influential variables retained by the combined model (i.e. CHH, resource trees in shrub layer and ground vegetation cover) had almost identical effects on presence/pseudo absence as in the LiDAR and field model (Figure 12 and 11) . The clumpyness of the canopy class, however, was distinguished by two thresholds at the values 0.75 and 0.83. Each of these values indicated a considerable decline in the response curve. The bilberry cover reached its maximum positive effect on the response at 28%. The effect of fern cover on the response steadily increased until a positive function was fitted at 7.5%, after which it stayed constant. As in the field model, the variable basal branched trees had a threshold of 2. The Shannon's Diversity Index for the CHM classes 1 & 2 had more influence in the combined than in the LiDAR model. Its threshold lay at a value of 0.5, where the fitted function flipped from negative to positive. The effect of tall rowans was very similar to the detected effect in the field model: already a scarce occurrence of this resource increased the fitted function. The second threshold at a value of 4, as seen in the field model, was not apparent here. The effect of the ground vegetation distribution is similar to its effect in the field model.

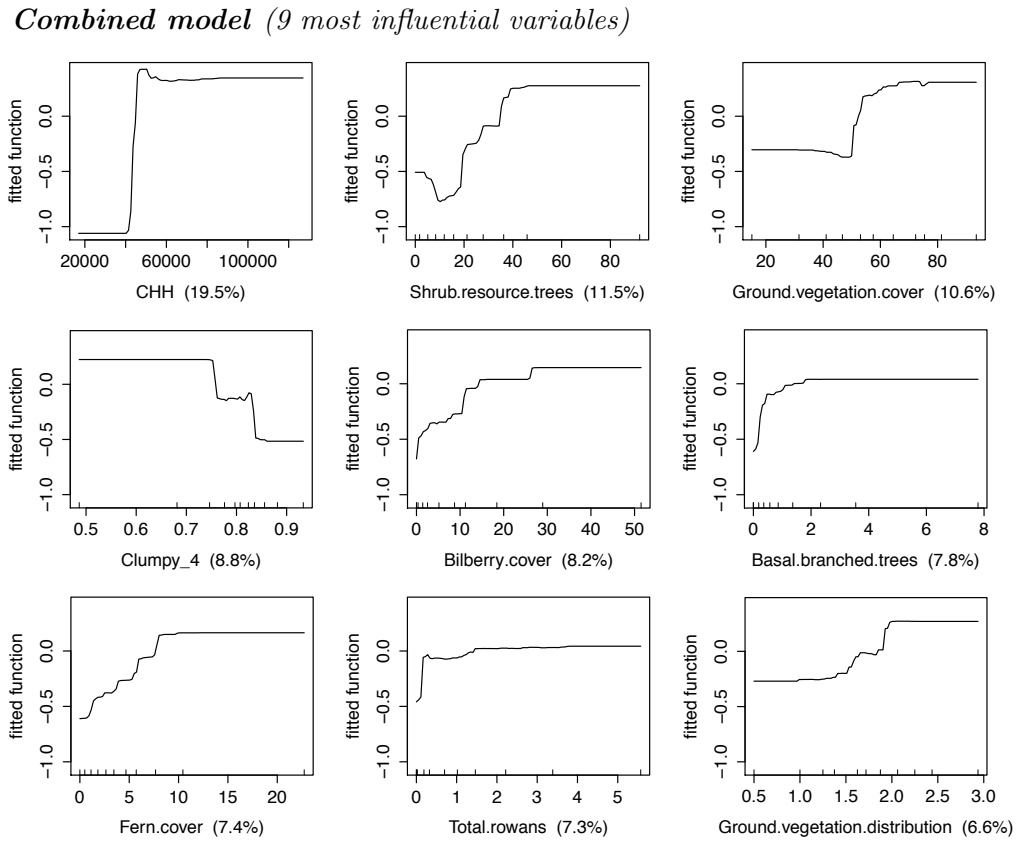


Figure 12: Partial dependence plots for 9 most influential variables retained by the final BRT combined model. The graphs show the effect of a particular variable on the response (i.e. presence/pseudo absence): increasing fitted function values indicate that hazel grouse responded positively, decreasing values the opposite. Values on the y-axis are uniformly scaled so that the relative importance of the variables can be visually compared [35]. The relative contribution of each variable to the total deviance explained is shown in parentheses.

Table 9: Variable interactions in the combined model.

Interacting Variables	Interaction size
Clumpy_4 - Resource trees in shrub layer	8.71
Ground vegetation distr. - Res. trees in shrub layer	6.92
Clumpy_4 - Fern cover	5.43
Bilberry cover - Fern cover	5.02
Ground vegetation distribution - $StdH_{80}$	3.88
Clumpy_4 - Bilberry cover	3.06

3.4 Variable interactions

The strongest interaction was detected between the clumpyness of the canopy class (Clumpy_4) and the cover of resource trees in the shrub layer, however, none of the identified interactions seem to dominate (Table 9). 4 of the 6 interactions include a combination of LiDAR and field derived variables, which makes it difficult for interpretation. It is assumed that interactions did not substantially influence model outcomes and are therefore not further dealt with in the discussion section.

3.5 Altitude and regional aspects

As indicated throughout the model results in Table 10, including the variable altitude combined with excluding the data for Central Eastern Alps led to improved model performances. In the field model, the inclusion of the variable altitude alone did not improve the model performance, but excluding the data from Central Eastern Alps considerably improved the model. The LiDAR model is also strongly influenced by altitude and the Central Easter Alps, rendering good results when both factors were considered. The same applied for the combined model, however, the effect of altitude was less strong than the effect of the Central Eastern Alps. Table 10 is for information only and the impact of altitude and the study region Central Eastern Alps will not be taken up in the discussion section.

Table 10: Optimal settings, performance and AUC values for best performing BRT models relating hazel grouse presence and pseudo absence to field and LiDAR based variables, as well as combination thereof. All models were run in three versions: including altitude, excluding the Central Eastern Alps (CEA), and both. Beside the model name and the number of variables retained by each model, the table indicates the number of trees fitted (nt), at which learning rate (lr) and tree complexity, the total deviance explained and AUC values based on the training data (D^2, AUC), the total deviance explained and AUC values based on 10-fold cross-validation (CV D^2, CV AUC). Indicated figures represent the mean of 5 repeated runs of each model. The bag fraction is 0.75 (2.4.2).

Model	**No. variables**	**nt**	**lr**	**tc**	**D^2**	**AUC**	**CV D^2**	**CV AUC**
Field incl. alt.	7	1600	0.003	1	37.73	0.897	20.03	0.783
Field excl. CEA	12	1960	0.003	1	53.62	0.956	26.34	0.828
Field excl. CEA, incl. alt.	13	1550	0.003	1	52.25	0.951	28.02	0.837
LiDAR incl. alt.	4	1830	0.003	1	36.05	0.880	22.12	0.806
LiDAR excl. CEA	3	890	0.003	1	33.54	0.887	20.63	0.819
LiDAR excl. CEA, incl. alt.	4	1120	0.003	1	44.51	0.909	29.07	0.855
Combined incl. alt.	16	4840	<0.001	4	54.86	0.982	21.33	0.814
Combined excl. CEA	16	5380	<0.001	4	65.19	0.995	27.73	0.837
Combined excl. CEA, incl. alt.	17	5480	<0.001	4	66.72	0.994	29.41	0.848

4 Discussion

The assessment of hazel grouse habitat included a sample based field survey as well as an area-wide analysis of LiDAR based forest structural diversity. This twofold approach allows not only for a comprehensive discussion of habitat requirements but also of methodological implications and limitations deriving from the sampling design and spatial resolution of the species data. As further outlined below, LiDAR is able to overcome some of the limitations associated with the field sampling design and provided valuable insights into the forest structural diversity across different mountain regions.

Returning to the research questions, subsection 4.3 is based on the results obtained by the field model (Table 8 on page 28) and provides answers as to how hazel grouse presence relate to forest composition and structure. Further, 4.3 incorporates knowledge gained from the threshold values that are listed in Figure 11 on page 30. The field assessment revealed that hazel grouse presence was strongly related to characteristics of the field layer, the availability of food plants and the abundance of basal branched trees. Thresholds were derived for all variables, however, not all of the thresholds were as clear as for the variables ground vegetation cover, basal branched trees or total rowans.

Several LiDAR variables related to both, the horizontal and vertical structures were successfully used to predict hazel grouse habitat suitability. Whereby variables describing the horizontal forest structure, e.g. canopy height heterogeneity, were more powerful in explaining habitat suitability than variables of the vertical structure, such as the standard deviations among the height band percentiles. The ability of LiDAR to analyse the structural diversity of mountain forests in relation to hazel grouse habitat suitability is discussed in 4.3.1 and 4.4, which are mainly based on the results presented in Table 8.

Based on the model evaluation criteria listed in Table 7 on page 27, it becomes evident in 4.2 that field and LiDAR assessments compared very well to each other in terms of model performance. However, the combined approach rendered superior results by incorporating the strengths of the field assessment (i.e. shrub composition and field layer) and of the LiDAR assessment (i.e. area-wide forest structure).

4.1 Study design

The species data used in this study has a resolution of $1km^2$. The field based habitat survey involved a sampling design with plot based measurements, hence, only a small fraction (1/70) of each $1km^2$ cell was sampled. The LiDAR based assessments, on the other hand, covered the whole area of each cell. This methodological framework raises some important implications.

Most previous assessments of hazel grouse habitat are based on measurements at the

core of used habitat or territories, implying a finer than 1km^2 grain size and thus increasing the resolution in the data [79, 65, 70, 78]. A higher resolution in the data should be reflected in more accurate modelling results and higher thresholds in limiting resources. But an increase of resolution often takes place at the cost of the spatial extent of the study area. With a grain size of 1km^2, however, it is more than likely that one will find a highly variable mountain forest environment, encompassing a broad range of stands differing in habitat quality, in presence as well as in pseudo absence cells. This introduces a considerable amount of heterogeneity to the data and is driven by factors such as the topography, local changes of site condition, local disturbances (windthrow, snow movements and breakage, insects, diseases, fire), and forest management practices. However, the availability of area based species data made it possible to expand the study area of this analysis to cover much of the variability of hazel grouse habitat in Swiss mountain forests.

Another aspect concerns the spatial coverage and accuracy of the field survey compared to LiDAR based assessments. Neither the field, nor the combined model retained any structural variables measured in the field, but structural variables based on LiDAR contributed substantially to the combined model. This implies that the plot sampling design was probably not detailed enough to represent relevant forest structures or that the size of the field plots was insufficient to quantify the relevant variation of forest structures. Due to their ambiguity, forest structures are also difficult to quantify in the field, especially if several people are involved in the sampling. LiDAR derived variables therefore provide valuable alternatives to field based assessments of forest structure, especially when large areas need to be sampled. As discussed below, LiDAR variables are likely to serve as proxies for several habitat characteristics relating to forest composition. However, the model results suggest that field based assessments were more accurate in quantifying the resource related forest composition as well as aspects of the ground vegetation, both being major characteristics of hazel grouse habitat.

4.2 Model comparison and assessment

All three final models were able to explain the observed data well with the combined model rendering the best results and the field model performing slightly better than the LiDAR model. According to the AUC interpretation by Hosmer and Lemenshow [40], the field and the LiDAR model achieved an excellent discrimination power, the combined model even an outstanding value of 0.977. This provides evidence that the field and LiDAR variables complemented each other very well, which is in line with other studies that found superior results when remotely sensed and field based data were combined [72, 26].

Despite the excellent results in explaining the observed data, model outcomes based on cross validated values decreased considerably. Cross validation accounts for over-optimism

and allows for evaluating the predictive performance of species distribution models [33]. Nevertheless, the cross validated combined model could still explain a substantial amount of deviance (23.5%) and achieved a good discrimination ability (AUC 0.804). It has to be noted that this study did not primarily aim at establishing a predictive model, but rather tried to assess habitat parameters across a large spatial extent, thus covering a wide range of prevailing conditions in Swiss mountain forests.

Beside the promising outcome, several aspects of boosted regression trees as a modelling technique facilitated and improved the interpretation of the results. Apart from the independence of predictor variable types and the ability to model nonlinearities, boosted regression trees allowed for deriving thresholds, which can serve as practical guidelines for forest management recommendations. Furthermore, interactions between variables were automatically identified and fitted, which in this case revealed that variable interactions were weak and did not strongly influence the outcome, a fact that is not easily detected by other modelling techniques.

4.3 Hazel grouse habitat characteristics in mountain forests

Various authors have studied the habitat requirements of hazel grouse in Switzerland [15, 52, 92, 78, 79] and elsewhere in Europe [27, 89, 6, 2, 70]. The results of this study are in line with these previous findings and confirm the importance of several habitat aspects such as a well developed field layer, an understory with deciduous softwoods as well as high structural diversity within inhabited forests.

Variables related to the field layer (i.e. ground vegetation cover, distribution and composition) characterised optimal habitat elements very well. Apart from hiding opportunities, a rich field layer can provide food in form of berries or buds, or indirectly in form of insects, and thus covers crucial habitat requirements for this ground nesting bird. However, it should be noted that very dense ground vegetation can also hamper movement and the ability of the hazel grouse to monitor its surroundings, which impairs the habitat quality [27]. This impairment was reflected by the fact that a ground vegetation cover higher than the threshold of 50 to 60% could not further improve the habitat quality. Among the variables related to the field layer, its distribution might be one of the least studied indicators. Hazel grouse showed a strong preference towards an evenly distributed ground vegetation. Together with the threshold analysis, this adds to the picture of an optimal field layer profile consisting of a small-scale mosaic of dense ground vegetation patches in combination with the occurrence of bilberry (Figure 13). To a large extent, such a field layer profile is determined by sufficient light conditions. Basal branched trees (which were predominantly spruce (*Picea abies*)) also depend on the availability of light and occur for example in forest stands with a sparse canopy closure or at forest edges.

Figure 13: Dense ground vegetation with patches of dwarf-shrubs consisting of bilberry (*Vaccinium myrthillus)* and juniper (*Juniperus communis*). Several basal branched spruce can also be seen.

These trees provide excellent hiding and resting opportunities and it is thus not surprising that the co-occurrence of a rich field layer and basal branched trees considerably improve the habitat quality for hazel grouse.

The availability and accessibility of food plants is known to considerably influence the occurrence of hazel grouse [15, 78, 79]. As in previous studies (e.g.[79]), tall rowan trees were identified as crucial habitat elements. Especially during wintertime, when other potential resources are not accessible due to the snow cover, rowan buds and berries can serve as an important food source [52, 50, 78]. The threshold analysis of this variable revealed that four tall rowans (>3m) per hectare substantially improved the habitat quality. Schäublin and Bollmann [79] found an average number of 23 tall rowans (>5m) per hectare in used stands. This discrepancy most likely derived from the implications of the different resolutions in the species data as discussed in 4.1. However, the availability of rowans also differs regionally and lower densities of rowans can partially be compensated by other resource related habitat elements (e.g. higher availabilities of other deciduous softwoods).

The occurrence of resource trees in the shrub layer is important for two reasons: first, the limited visibility in shrubs provide protection from predators, and second, the birds can feed on catkins, buds and berries. With the inconspicuous appearance and agile flying capabilities, hazel grouse are very well adapted to this key habitat characteristic. Although [79] and [65] found that the shrub cover was an important habitat variable, the

Figure 14: Fructifying rowan trees and other deciduous softwoods growing in a former Lothar windthrow area.

results of this analysis suggest that the composition of the shrub layer with respect to available resources is most relevant. This is in line with previous studies reporting that the occurrence of deciduous softwoods in the understory is a key characteristic of hazel grouse habitat [2, 70, 78].

Rowans and other deciduous softwoods are typical colonisers of early successional stages, which are naturally induced by windthrow events, insect calamities, avalanches or other disturbances that change the microclimatic conditions and increase the availability of light on the ground (Figure 14). Therefore, it is likely that areas affected by past windthrow events like Vivian (1990) and Lothar (1999), together with subsequent bark beetle infestations, provided additional primary habitat for hazel grouse (see also [79]). However, early successional stages also emerge following forest management interventions such as group selection cuttings. Among other measures like promoting deciduous softwoods, leaving clearings to natural succession is an appropriate forest management practice to improve the habitat quality for hazel grouse [15, 79].

4.3.1 Structural diversity

The horizontal and vertical diversity in forest structure is an important driver of forest biodiversity in general [49, 62] and hazel grouse habitat quality in particular [79, 65, 70]. Structural elements and patterns are characterised by differing light conditions, which

makes a light based method such as LiDAR suitable for detecting structural changes within a forest. This is most clearly expressed by the important model contribution of the LiDAR variable canopy height heterogeneity. A forest with stands varying in canopy heights (as defined in 2.3.2.3) is characterised by alternating light conditions in the understory and field layer, which in turn promotes the ground vegetation, rejuvenation of shrubs, single trees or groups thereof and a patchy distribution of these elements. Furthermore, varying canopy heights also reflect transitions between different successional stages or linear elements within the forest, like brooks, forest edges or forest roads. These linear ecotones are particularly interesting because they change the local light conditions in a way that allows for the emergence of light demanding resource shrubs and trees, which are frequently used by hazel grouse [91, 79]. In the immediate vicinity of brooks (e.g. eroded banks) or forest roads, it is also likely that hazel grouse can find places for sand bathing and grit ingestion [6]. However, it has to be kept in mind that the variable canopy height heterogeneity is not an indicator for abrupt transitions between forest and non-forest. In fact, its strength lies in quantifying smooth transitions from forest to non-forest areas. Due to the patchy occurrence of deciduous shrubs and tree rejuvenation, which provide foraging opportunities and sufficient cover over small distances, smooth transition zones are potential habitats for hazel grouse. Figure 15 illustrates how such transition zones are captured in the classified canopy height model. During the past decades, these transition zones have spread in mountain regions of Switzerland mainly due to former farmland being left unmanaged and subsequent ingrowth [25, 81]. It is likely that additional hazel grouse habitat has developed from this land-use change.

Single layered, mature forest stands with a sparsely developed or missing understory are avoided by hazel grouse because they provide neither cover nor resources [91, 70]. An increasing clumpyness of the canopy class (see 2.3.2.5) indicates such stands and confirms that if they expand over large areas, the habitat quality is significantly reduced. Further, hazel grouse favoured areas with smaller standard deviations within the vertical forest structure. On a first glance, this seems counterintuitive because a high standard deviation within the forest vegetation could be interpreted as structural richness and therefore positive for hazel grouse. But here the situation slightly differs in a sense that a smaller standard deviation among a particular band percentile (see 2.3.2.3) reflect a more evenly distributed, multi-layered vertical forest structure, in contrast to a vertical profile that is dominated by a single or two layered structure resulting in high standard deviations. Figure 16 illustrates two contrasting vertical profiles. The vegetation heights of the multi-storied forest (top) vary quite evenly between 2 and 18 meters. The single layered profile (bottom), on the other hand, shows a main distribution of vegetation heights between 25 and 40, however, some points also reach further down. Apart from the higher maximum

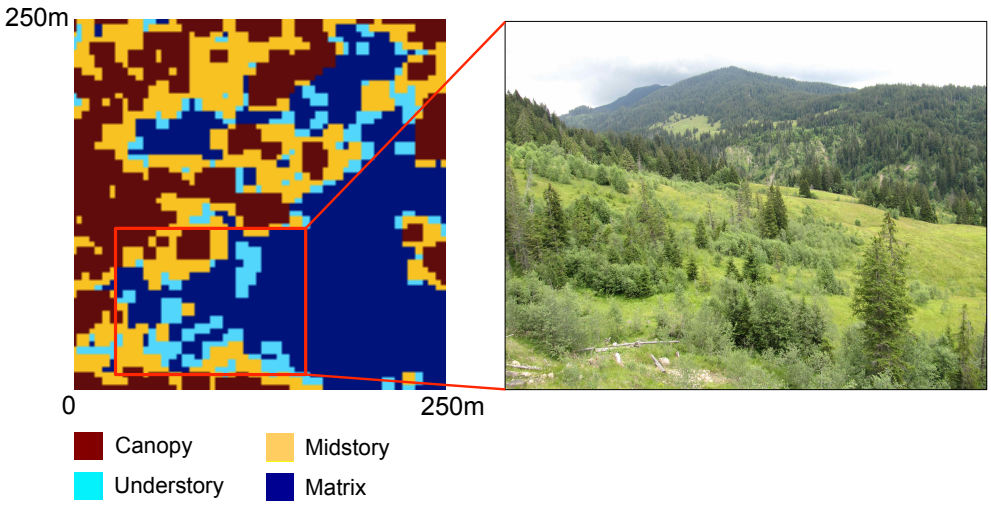


Figure 15: Extract of classified canopy height model with a pixel size of 4m showing continuous transitions from forest to open land due to ingrowth (abandonment of farming) (left). The transition zones are characterised by a patchy occurrence of deciduous pioneer woods, spruce rejuvenation and alpine meadows (right).

height in the single layered profile, the small amount of laser returns in the middle together with some returns close to the ground lead to an increased data range and result in a higher standard deviation, especially among the middle and lower height band percentiles.

4.4 LiDAR based habitat assessments of forest species

The strength of LiDAR to accurately capture relevant characteristics of forest structure has made it a promising tool for habitat and biodiversity assessments [5, 86, 11, 14]. Goetz and colleagues [29] found that LiDAR derived variables characterising the vertical distribution and complexity of canopy elements in an eastern temperate forest in the USA performed well in terms of predicting bird species richness (avg. 30-40% of variance explained). Their analysis base on full waveform LiDAR data recording a larger number of laser return signals and thus reflecting vegetation structure more precisely than first and last return LiDAR data as applied in the study at hand. Also, the authors involve abiotic factors like ground elevation and LiDAR return intensity related variables, both being not included or not available in this study. Interestingly, combining LiDAR variables with potential optical predictors like the Normalised Differenced Vegetation Index (NDVI) did not considerably improve their results, which underlines the strength of LiDAR to quantify relevant habitat characteristics. A very similar approach by partially the same authors resulted in an accurate prediction of breeding habitat of a Neotropical migrant bird (up to 47% variance explained) [28]. Lesak et al. [60] modelled forest songbird species richness in an upland forest in Wisconsin (USA) and were able to explain their observed

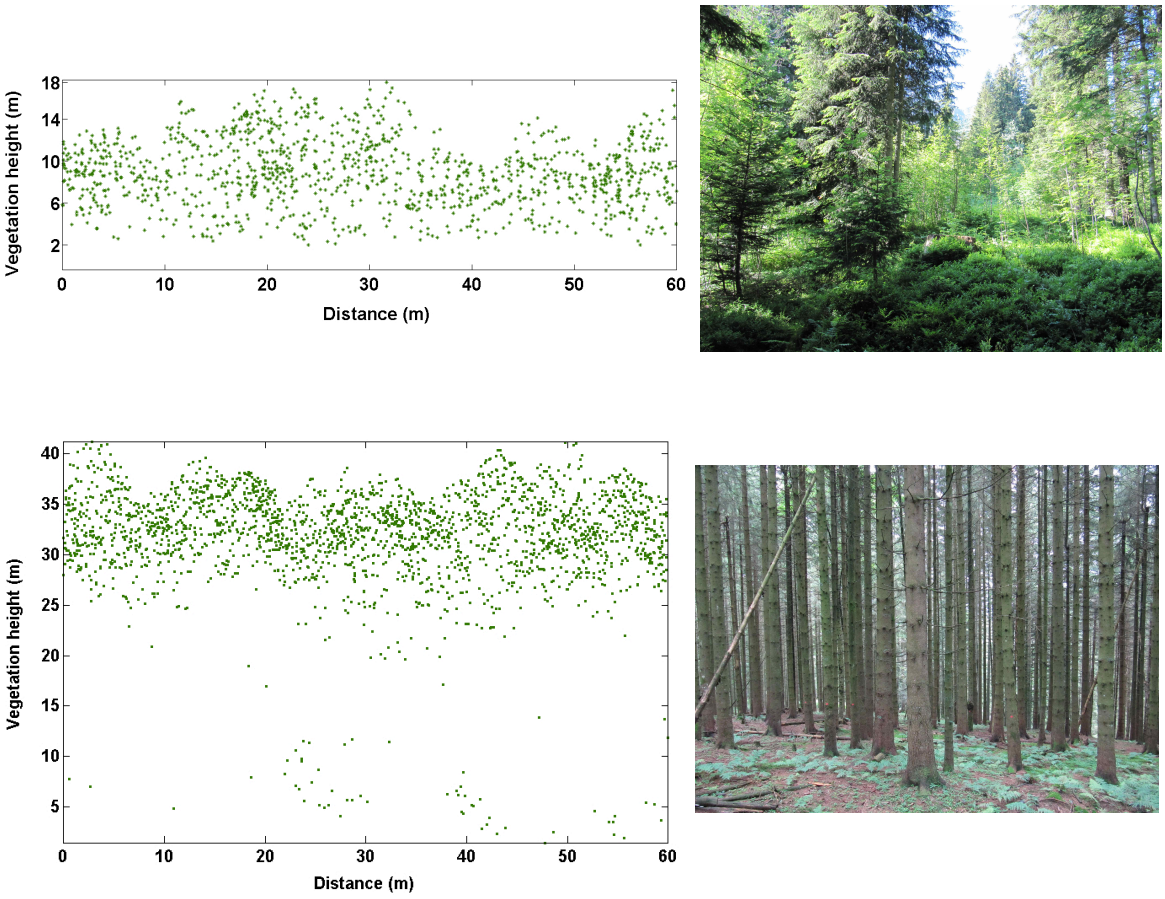


Figure 16: Vertical distribution of vegetation point heights (laser returns) in a multi-layered (top) and single-layered forest stand (bottom); both being illustrated by a real life example on the right.

data moderately well, with explained variances from 5 to 21%. Müller et al [72] analysed the potential abundance and assemblages of forest passerines in Germany and found that the predictive power of LiDAR variables was, with a few exceptions, superior to that of field measurements and visual interpretation of areal photographs. The explained variance based on predicted data for single species ranged from < 5 to 40%, with the mean canopy height being the most important variable. Hinsley et al [38, 39] compared LiDAR derived canopy heights and cover with bird species distributions in a coppice forest in Britain. They found relatively high correlations and showed that the canopy height could serve as surrogate of woodland structure and thus be used to predict woodland bird composition and distribution in lowland forests in Britain. LiDAR derived variables were also able to improve habitat models for capercaillie (*Tetrao urogallus*) in a montane forest in Switzerland [32]. Furthermore, LiDAR variables performed well in predicting spider distributions [85] and beetle assemblages [71] in mountainous forests in south-eastern Germany.

An important common characteristic among these cited studies is that they involve local scale analysis based on point observations with the implications as discussed in 4.1. Nevertheless, the result of the LiDAR model in this study, with 34% explained deviance in the observed data, compares well with previous studies. This underlines the remarkable potential of LiDAR derived metrics for assessing habitat characteristics across broad spatial scales, even in a highly heterogeneous mountain forest environment.

5 Conclusions

5.1 Achievements

This study includes an assessment of habitat requirements of the endangered hazel grouse over a broad range of mountain forest conditions, encompassing three biogeographic regions in Switzerland. None of the previous studies of hazel grouse habitat in Central Europe were carried out over such a large spatial extent. Thus, a major achievement of this thesis is to show that the current knowledge of the habitat requirements can be applied to a large study area including a high variability of hazel grouse habitat.

The field survey involved a sample plot based assessment of habitat elements related to forest composition and structure and revealed that the sampling design was able to provide valuable information with respect to the field layer as well as the composition of the forest vegetation. However, relevant forest structures were captured more accurately by the area-wide assessment based on LiDAR data.

The LiDAR based habitat assessment is an important element of this thesis, because scientific knowledge about LiDAR based habitat analysis is still relatively scarce, although the potential is expected to be high. Therefore, it was essential to develop a method that incorporated current advances in terms of LiDAR data processing with ecological knowledge. This was done by writing a MATLAB routine using functions (provided by Felix Morsdorf) that allowed for an efficient transformation of LiDAR point data into ecologically relevant parameters of forest structures. Parameters for the vertical structure of forests were directly computed within the MATLAB routine. In order to derive parameters characterising the horizontal forest structure, intermediate results obtained in MATLAB were further analysed in FRAGSTATS, a tool to analyse spatial patterns and landscape structures. The twofold approach of processing LiDAR data resulted in a statistical model that performed almost as good as the model based on the field assessment. This is remarkable considering that hazel grouse is a distinctive food specialist and that compositional aspects of the forest vegetation cannot directly be measured with the LiDAR data at hand.

5.2 Insights and outlook

Given the large spatial resolution of the species data in this study, the main advance herein is the ability of LiDAR to detect relevant forest structures area-wide, thus overcoming the difficulties associated with sample based field assessments. However, the field survey confirmed that hazel grouse depends very much on the structure and composition of the field and shrub layer. These habitat characteristics were difficult to quantify with the LiDAR data at hand, because laser returns that reached as far down as the field layer are difficult to distinguish from the terrain and the chances to detect understory returns is

reduced when the cover of higher stories increases [31]. Further, LiDAR cannot directly measure the composition of the field and shrub layer in terms of resource plant species. But structural features captured in LiDAR, such as the canopy height heterogeneity, may serve as indicators for thermal and light conditions on the forest floor and understory, thus serving as proxies that indirectly reflect aspects of plant species composition. Based on promising results obtained in recent studies, e.g. [64] and [28], it is assumed that part of these shortcomings could be compensated by incorporating multiple return or full waveform LiDAR data that depict a more detailed picture of the vertical forest structure. Calibrated LiDAR intensity data, as applied by e.g. [68] and [55], was also found to be sensitive to the understory and ground-surface composition, thus holding potential to improve LiDAR based assessments of wildlife habitat suitability. Further, [68] suggests that the use of multi-spectral canopy LiDAR could be useful for tree species distinction and to identify different strata of vegetation.

The ability of LiDAR to assess the habitat availability and quality area-wide renders several advantages for a strengthened integration of conservation aspects into forest management. In a fragmented landscape like the Swiss mountain forests, suitable hazel grouse habitat is dispersed in more or less isolated patches differing in size and quality. Therefore, knowledge about the habitat connectivity is crucial for conserving this sedentary bird species with limited dispersal abilities [82, 27]. The ability to identify suitable habitat in the form of isolated patches or corridors that connect two habitat patches makes LiDAR a valuable tool to assess the structural connectivity within a landscape [83]. However, LiDAR derived metrics also assess the quality of the forest and landscape matrix and therefore provide insights about the functional connectivity. Thus, LiDAR provides adequate means to integrate aspect relevant to the conservation of viable populations into forest planning instruments, e.g. cantonal forest development plans.

In the light of emerging conservation strategies such as the biodiversity strategy proposed by the Swiss Federal Office for the Environment (FOEN) [22], knowledge about the spatial distribution and the degree of forest structural diversity will become valuable. LiDAR based methods, as proposed here, could help to provide this knowledge and could be used to identify hotspots of forest structural richness. Apart from providing spatially extensive information about the vegetation structure, assessments based on LiDAR data are less labour and cost intensive than field based assessments. However, specific software and expert know-how is required to transform the data into user relevant information [67]. Last but not least, multi-temporal data acquisition and photogrammetric updates of LiDAR derived canopy height models would allow for change detection and could be used to monitor and control the effects of conservation related to forest management practices.

References

[1] *R Foundation for Statistical Computing.* R: a language and environment for statistical computing. Package version 2.13.2. Vienna, Austria. `http://www.R-project.org`, 2011.

[2] J. Aberg, J.E. Swenson, and P. Angelstam. The habitat requirements of hazel grouse (*Bonasa bonasia*) in managed boreal forest and applicability of forest stand descriptions as a tool to identify suitable patches. *Forest Ecology and Management*, 175(1-3):437–444, 2003.

[3] E.P. Baltsavias. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing Remote Sensing*, 54(2-3):199–214, 1999.

[4] J. Bengtsson, S.G. Nilsson, A. Franc, and P. Menozzi. Biodiversity, disturbances, ecosystem function and management of European forests. *Forest Ecology and Management*, 132(1):39–50, 2000.

[5] K.M. Bergen, S.J. Goetz, R.O. Dubayah, G.M. Henebry, C.T. Hunsaker, M.L. Imhoff, R.F. Nelson, G.G. Parker, and V.C. Radeloff. Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research-Biogeosciences*, 114, 2009.

[6] H.-H Bergmann, S. Klaus, Müller F., Scherzinger W., J.E. Swenson, and J. Wiesner. *Die Haselhühner - 4. überarbeitete Auflage.* Die neue Brehm-Bücherei Bd.77. Westarp Wissenschaften, Magdeburg, 1996.

[7] M. Blattner. Der Arealschwund des Haselhuhns *Bonasa bonasia* in der Nordwestschweiz. *Der Ornithologische Beobachter*, 95(1):11–38, 1998.

[8] K. Bollmann, A. Bergamini, B. Senn-Irlet, M. Nobis, P. Duelli, and Scheidegger C. Konzepte, Instrumente und Herausforderungen bei der Förderung der Biodiversität im Wald. *Schweizerische Zeitschrift für Forstwesen*, 160:53–67, 2009.

[9] K. Bollmann, V. Braunisch, and R. Arlettaz. Predicted effects of climate change on indicator species of structural and biological diversity in mountain forests: towards adaptive forest managment in the face of environmental uncertainty. Research proposal, unpublished.

[10] K. Bollmann, V. Keller, W. Müller, and N. Zbinden. Prioritäre Vogelarten für Artenförderungsprogramme in der Schweiz. *Der Ornithologische Beobachter*, 99:301–320, 2002.

[11] R.B. Bradbury, R.A. Hill, D.C. Mason, S.A. Hinsley, J.D. Wilson, H. Balzter, G.Q.A. Anderson, M.J. Whittingham, I.J. Davenport, and P.E. Bellamy. Modelling relationships between birds and vegetation structure using airborne lidar data: a review with case studies from agricultural and woodland environments. *Ibis*, 147(3):443–452, 2005.

[12] U.-B. Brändli (Eds.). *Schweizerisches Landesforstinventar. Ergebnisse der dritten Erhebung 2004-2006*. Birmensdorf, Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft WSL. Bern, Bundesamt für Umwelt, BAFU, 2010.

[13] P.M. Buston and J. Elith. Determinants of reproductive success in dominant pairs of clownfish: a boosted regression tree analysis. *Journal of Animal Ecology*, 80(3):528–538, 2011.

[14] R.M. Clawges, K.T. Vierling, L.A. Vierling, and E. Rowell. The use of airborne lidar to assess avian species diversity, density, and occurrence in a pine/aspen forest. *Remote Sensing of Environment*, 112(5):2064–2073, 2008.

[15] K. Eiberle and N. Koch. Die Bedeutung der Waldstruktur für die Erhaltung des Haselhuhns (*Tetrastes bonasia L.*). *Schweizerische Zeitschrift für Forstwesen*, 126(12):876–887, 1975.

[16] J. Elith and C.H. Graham. Do they? How do they? Why do they differ? on finding reasons for differing performances of species distribution models. *Ecography*, 32(1):66–77, 2009.

[17] J. Elith, C.H. Graham, R.P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J.M. Overton, A.T. Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberon, S. Williams, M.S. Wisz, and N.E. Zimmermann. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2):129–151, 2006.

[18] J. Elith and J.R. Leathwick. *Boosted Regression Trees for ecological modeling.* `http://cran.r-project.org/web/packages/dismo/vignettes/brt.pdf` (24.11.2011), October 2011.

[19] J. Elith, J.R. Leathwick, and T. Hastie. A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4):802–813, 2008.

[20] A.H. Fielding and J.F. Bell. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1):38–49, 1997.

[21] A.H. Fielding and P.F. Haworth. Testing the generality of bird-habitat models. *Conservation Biology*, 9(6):1466–1481, 1995.

[22] Federal Office for the Environment FOEN. *Strategie Biodiversität Schweiz.* `http://www.bafu.admin.ch/biodiversitaet/10372/10395/index.html` (11.01.2012).

[23] J.F. Franklin, T.A. Spies, R. Van Pelt, A.B. Carey, D.A. Thornburgh, D.R. Berg, D.B. Lindenmayer, M.E. Harmon, W.S. Keeton, D.C. Shaw, K. Bible, and J.Q. Chen. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. *Forest Ecology and Management*, 155(1-3):399–423, 2002.

[24] J.H. Friedman. Stochastic gradient boosting. *Computational Statistics and Data Analysis*, 38(4):367–378, 2002.

[25] J. Gehrig-Fasel, A. Guisan, and N.E. Zimmermann. Tree line shifts in the Swiss Alps: Climate change or land abandonment? *Journal of Vegetation Science*, 18(4):571–582, 2007.

[26] L.A. Gibson, B.A. Wilson, D.M. Cahill, and J. Hill. Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *Journal of Applied Ecology*, 41(2):213–223, 2004.

[27] U.N. Glutz von Blotzheim (eds), K.M. Bauer, and E. Bezzel. *Handbuch der Vögel Mitteleuropas*, volume 5: Galliformes und Gruiformes. Akademische Verlagsgesellschaft, Frankfurt am Main, 1973.

[28] S. J. Goetz, D. Steinberg, M. G. Betts, R. T. Holmes, P. J. Doran, R. Dubayah, and M. Hofton. Lidar remote sensing variables predict breeding habitat of a Neotropical migrant bird. *Ecology*, 91(6):1569–1576, 2010.

[29] S.J. Goetz, D.l Steinberg, R.O. Dubayah, and B. Blair. Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sensing of Environment*, 108(3):254–263, 2007.

[30] Y. Gonseth, T. Wohlgemuth, B. Sansonnens, and A. Buttler. *Die biogeographischen Regionen der Schweiz. Erläuterungen und Einteilungsstandard.* Umwelt Materialien Nr. 137 Bundesamt für Umwelt, Wald und Landschaft Bern., 2001.

[31] N.R. Goodwin, N.C. Coops, C.W. Bater, and S.E. Gergel. Assessment of sub-canopy structure in a complex coniferous forest. *Proceedings of the ISPR Workshop "Laser Scanning 2007 and SilviLaser 2007", Espoo, September 12-14, 2007, Finland, Vol. XXXVI.*, ISSN:1682-1777, P3/W52:169–172, 2007.

[32] R.F. Graf, L. Mathys, and K. Bollmann. Habitat assessment for forest dwelling species using LiDAR remote sensing: Capercaillie in the Alps. *Forest Ecology and Management*, 257(1):160–167, 2009.

[33] A. Guisan and N.E. Zimmermann. Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2-3):147–186, 2000.

[34] L. Hannah, J. L. Carr, and A. Landerani. Human disturbance and natural habitat - a biome level analysis of a global data set. *Biodiversity and Conservation*, 4(2):128–155, 1995.

[35] T. Hastie, R. Tibshirani, and J.H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics. Springer Science and Business Media, Inc., 233 Spring St., New York, USA, 2001.

[36] R.J. Hijmans, S.J. Phillips, J.R. Leathwick, and J. Elith. Species distribution modeling (dismo). package version 0.7-8. `http://cran.r-project.org/web/packages/dismo/index.html`. 2011.

[37] R.A. Hill, S.A. Hinsley, D.L.A. Gaveau, and P.E. Bellamy. Predicting habitat quality for Great Tits (*Parus major*) with airborne laser scanning data. *International Journal of Remote Sensing*, 25(22):4851–4855, 2004.

[38] S.A. Hinsley, R.A. Hill, R.J. Fuller, P.E. Bellamy, and P. Rothery. Bird species distributions across woodland canopy structure gradients. *Community Ecology*, 10(1):99–110, 2009.

[39] S.A. Hinsley, R.A. Hill, D.L.A. Gaveau, and P.E. Bellamy. Quantifying woodland structure and habitat quality for birds using airborne laser scanning. *Functional Ecology*, 16(6):851–857, 2002.

[40] D.W. Hosmer and S. Lemeshow. *Applied Logistic Regression*. John Wiley and Sons. Second edition, New York, USA, 2000.

[41] P. Hyde, R. Dubayah, B. Peterson, J.B. Blair, M. Hofton, C. Hunsaker, R. Knox, and W. Walker. Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sensing of Environment*, 96(3-4):427–437, 2005.

[42] P. Hyde, R. Dubayah, W. Walker, J.B. Blair, M. Hofton, and C. Hunsaker. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment*, 102(1-2):63–73, 2006.

[43] J. Hyyppä, H. Hyyppä, D. Leckie, F. Gougeon, X. Yu, and M. Maltamo. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing*, 29(5):1339–1366, 2008.

[44] M.P. Indermühle. Struktur-, Alters- und Zuwachsuntersuchungen in einem Fichten-Plenterwald der subalpinen Stufe (Sphagno-Piceetum calamagrostietosum villosae). *Beiheft zu den Zeitschriften des Schweizerischen Forstvereins*, (60):97, 1978.

[45] J. Jones. Habitat selection studies in avian ecology: A critical review. *Auk*, 118(2):557–562, 2001.

[46] V. Keller and K. Bollmann. Für welche Vogelarten trägt die Schweiz eine besondere Verantwortung? *Der Ornithologische Beobachter*, 98:323–340, 2001.

[47] V. Keller, A. Gerber, H. Schmid, P. Volet, and N. Zbinden. *Rote Liste der gefährdeten Brutvogelarten der Schweiz.* Bundesamt für Umwelt, Wald und Landschaft, Bern, und Schweizerische Vogelwarte, Sempach. Umwelt-Vollzug Nr. 1019, 2010.

[48] M. Keller (eds.). *Schweizerisches Landesforstinventar. Anleitung für die Feldaufnahmen der Erhebung 2004-2007.* Birmensdorf. Eidg. Forschungsanstalt WSL, 2005.

[49] J.P. Kimmins. *Forest Ecology - A Foundation for Sustainable Forest Management and Environmental Ethics in Forestry.* Prentice Hall, Upper Saddle River, New Jersey, third edition, 2004.

[50] S. Klaus. Effects of forestry on grouse populations - case-studies from the Thuringian and Bohemian forests, Central-Europe. *Ornis Scandinavica*, 22(3):218–223, 1991.

[51] S. Klaus, J. Martens, and Sun Y-.H. Andreev A.V. *Bonasa bonasia (Linnaeus, 1758).* Atlas Verbr Palaearkt Vögel 20:1-15, 2003.

[52] N. Koch. Hasel- und Auerhuhn an der Hohen Rone (Kanton Zug, Schweiz). *Schweizerische Zeitschrift für Forstwesen*, 129(11):897–933, 1978.

[53] B. Koetz, F. Morsdorf, G. Sun, K.J. Ranson, K.I. Itten, and B. Allgower. Inversion of a Lidar Waveform Model for Forest Biophysical Parameter Estimation. *IEEE Geoscience and Remote Sensing Letters*, 3:49–53, 2006.

[54] B. Koetz, F. Morsdorf, S. van der Linden, T. Curt, and B. Allgöwer. Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data. *Forest Ecology and Management*, 256(3):263–271, 2008.

[55] I.S. Korpela. Mapping of understory lichens with airborne discrete-return LiDAR data. *Remote Sensing of Environment*, 112(10):3891–3897, 2008.

[56] T.-B. Larsson. *Biodiversity Evaluation Tools for European forests*. Ecological Bulletins 50: 000-000. 2001.

[57] J.R. Leathwick, J. Elith, M.P. Francis, T. Hastie, and P. Taylor. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Marine Ecology-Progress Series*, 321:267–281, 2006.

[58] J. Leclere, T. Oberdorff, J. Belliard, and F. Leprieur. A comparison of modeling techniques to predict juvenile 0+fish species occurrences in a large river system. *Ecological Informatics*, 6(5):276–285, 2011.

[59] M.A. Lefsky, W.B. Cohen, G.G. Parker, and D.J. Harding. Lidar remote sensing for ecosystem studies. *Bioscience*, 52(1):19–30, 2002.

[60] A.A. Lesak, V.C. C. Radeloff, T.J. Hawbaker, A.M. Pidgeon, T. Gobakken, and K. Contrucci. Modeling forest songbird species richness using LiDAR-derived measures of forest structure. *Remote Sensing of Environment*, 115(11):2823–2835, 2011.

[61] S. A. Levin. The problem of pattern and scale in ecology. *Ecology*, 73(6):1943–1967, 1992.

[62] D. B. Lindenmayer, C. R. Margules, and D. B. Botkin. Indicators of biodiversity for ecologically sustainable forest management. *Conservation Biology*, 14(4):941–950, 2000.

[63] N. Maag, K. Bollmann, and V. Braunisch. Wald, Klima und Bergvögel. Anleitung für die Feldaufnahmen. unpublished manual, 2011.

[64] S. Martinuzzi, L.A. Vierling, W.A. Gould, M.J. Falkowski, J.S. Evans, A.T. Hudak, and K.T. Vierling. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Remote Sensing of Environment*, 113(12):2533–2546, 2009.

[65] L. Mathys, N.E. Zimmermann, N. Zbinden, and W. Suter. Identifying habitat suitability for hazel grouse *Bonasa bonasia* at the landscape scale. *Wildlife Biology*, 12(4):357–366, 2006.

[66] K. McGarigal, S.A. Cushman, M.C. Neel, and E. Ene. FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. Computer software program. Technical report, University of Massachusetts, Amherst. Available at: `http://www.umass.edu/landeco/research/fragstats/fragstats.html` (14.10.2011), 2002.

[67] F. Morsdorf. Erfassung struktureller Waldparameter mithilfe von flugzeuggetragenem Laserscanning. *Schweizerische Zeitschrift für Forstwesen*, 162(2):164–170, 2011.

[68] F. Morsdorf, A. Marell, B. Koetz, N. Cassagne, F. Pimont, E. Rigolot, and B. Allgoewer. Discrimination of vegetation strata in a multi-layered Mediterranean forest ecosystem using height and intensity information derived from airborne laser scanning. *Remote Sensing of Environment*, 114(7):1403–1415, 2010.

[69] F. Morsdorf, E. Meier, B. Kötz, K.I. Itten, M. Dobbertin, and B. Allgöwer. LiDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. *Remote Sensing of Environment*, 3(92):353–362, 2004.

[70] D. Mueller, B. Schroeder, and J. Mueller. Modelling habitat selection of the cryptic Hazel Grouse *Bonasa bonasia* in a montane forest. *Journal of Ornithology*, 150(4):717–732, 2009.

[71] J. Mueller and R. Brandl. Assessing biodiversity by remote sensing in mountainous terrain: the potential of LiDAR to predict forest beetle assemblages. *Journal of Applied Ecology*, 46(4):897–905, 2009.

[72] J. Muller, C. Moning, C. Bassler, M. Heurich, and R. Brandl. Using airborne laser scanning to model potential abundance and assemblages of forest passerines. *Basic and Applied Ecology*, 10(7):671–681, 2009.

[73] E. Naesset. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80(1):88–99, 2002.

[74] E. Naesset. Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser. *Remote Sensing of Environment*, 91(2):243–255, 2004.

[75] Federal Department of Topography (Swisstopo). *VECTOR25*. `http://www.swisstopo.admin.ch/internet/swisstopo/de/home/products/landscape/vector25.html` (13.01.2012).

[76] Y. Paillet, L. Berges, J. Hjalten, P. Odor, C. Avon, M. Bernhardt-Roemermann, R.-J. Bijlsma, L. De Bruyn, M. Fuhr, U. Grandin, R. Kanka, L. Lundin, S. Luque, T. Magura, S. Matesanz, I. Meszaros, M. Teresa Sebastia, W. Schmidt, T. Standovar, B. Tothmeresz, A. Uotila, F. Valladares, K. Vellak, and R. Virtanen. Biodiversity Differences between Managed and Unmanaged Forests: Meta-Analysis of Species Richness in Europe. *Conservation Biology*, 24(1):101–112, 2010.

[77] J.-M. Roberge and P. Angelstam. Indicator species among resident forest birds - a cross-regional evaluation in northern Europe. *Biological Conservation*, 130(1):134–147, 2006.

[78] S. Sachot, N. Perrin, and C. Neet. Winter habitat selection by two sympatric forest grouse in western Switzerland: implications for conservation. *Biological Conservation*, 112(3):373–382, 2003.

[79] S. Schaeublin and K. Bollmann. Winter habitat selection and conservation of Hazel Grouse (Bonasa bonasia) in mountain forests. *Journal of Ornithology*, 152(1):179–192, 2011.

[80] W. Schönenberger. Trends in mountain forest management in Switzerland. *Schweizerische Zeitschrift für Forstwesen*, 152:152–156, 2001.

[81] J. Stöcklin, A. Bosshard, G. Klaus, K. Rudmann-Maurer, and M. Fischer. *Landnutzung und biologische Vielfalt in den Alpen - Thematische Synthese zum Forschungsschwerpunkt 2 "Land- und Forstwirtschaft im alpinen Lebensraum" des Nationalen Forschungungsprogramms NFP 48 "Landschaften und Lebensräume der Alpen" des Schweizerischen Nationalfonds SNF.* vdf Hochschulverlag AG, Zürich, 2007.

[82] I. Storch. Conservation status and threats to grouse worldwide: an overview. *Wildlife Biology*, 6(4):195–204, 2000.

[83] L. Tischendorf and L. Fahring. On the usage and measurement of landscape connectivity. *Oikos*, 90(1):7–19, 2000.

[84] W. Turner, S. Spector, N. Gardiner, M. Fladeland, E. Sterling, and M. Steininger. Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution Evolution*, 18(6):306–314, 2003.

[85] K.T. Vierling, C. Baessler, R. Brandl, L.A. Vierling, I. Weiss, and J. Mueller. Spinning a laser web: predicting spider distributions using LiDAR. *Ecological Applications*, 21(2):577–588, 2011.

[86] K.T. Vierling, L.A. Vierling, W.A. Gould, S. Martinuzzi, and R.M. Clawges. Lidar: shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment*, 6(2):90–98, 2008.

[87] L.A. Vierling, S. Martinuzzi, G.P. Asner, J. Stoker, and B.R. Johnson. Lidar: providing structure. *Frontiers in Ecology and the Environment*, 9(5):261–262, 2011.

[88] W. Wagner, A. Ullrich, and C. Briese. *Der Laserstrahl und seine Interaktion mit der Erdoberfläche*, volume 4. 2003.

[89] J. Wiesner, H.-H. Bergmann, and S. Klaus. Siedlungsdichte und Habitatstruktur des Haselhuhns (*Bonasa bonasia*) im Waldgebiet von Bialowieza (Polen). *Journal für Ornithologie*, 118(1):1–20, 1977.

[90] O. Wildi and T. Wohlgemuth. Waldstatik - Walddynamik. *WSL Informationblatt Forschungsbereich Landschaft*, 69:1–4, 2007.

[91] N. Zbinden. Zur Ökologie des Haselhuhns (*Bonasa bonasia*) in den Buchenwäldern des Chasseral, Faltenjura. *Der Ornithologische Beobachter*, 76:169–214, 1979.

[92] N. Zbinden and M. Blattner. *Haselhuhn. Schweizer Brutvogelatlas - Verbreitung der Brutvögel in der Schweiz und im Fürstentum Lichtenstein 1993-1996.* Schweizerische Vogelwarte, Sempach, 1998.

[93] D.A. Zimble, D.L. Evans, G.C. Carlson, R.C. Parker, S.C. Grado, and P.D. Gerard. Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment*, 87(2-3):171–182, 2003.

Acknowledgements

Kurt Bollmann was the main adviser of this Master thesis. I am very grateful for his expert input throughout the working process and thank him for making it all possible. Felix Morsdorf's expertise was indispensable for the processing of the LiDAR data; his numerous spontaneous answers were not only helpful but also very motivating. I thank Ross Purves for representing this thesis at the Faculty and providing important methodological hints as well as valuable comments to the text. I extend my appreciation to all the people involved in the field assessments, especially to Lisa Bitterlin, who showed great endurance in the difficult terrain and all weather conditions. My cheerful Martina always supported me and we had great fun in the mountains.

A Statistical overview of field and LiDAR derived habitat variables

Table 11: Statistical overview of field and LiDAR derived habitat variables. Values are indicated for presence and pseudo absence cells separately. Significance (*Sig.*) is based on univariate generalised linear models (GLMs). Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Study area (S), Jura (J), Northern Prealps (P), Central Eastern Alps (C). The correlation of indicated variables is below 0.7 (based on Spearman's rank correlation coefficient), however, field and LiDAR variables were not tested against each other due to assumed independence.

	Presence Grids				*Absence grids*				*Sig.*			
Field variables	**Min**	**Max**	**Mean**	**Sd**	**Min**	**Max**	**Mean**	**Sd**	**S**	**J**	**P**	**C**
Altitude	804	2'020	1'409	208.3	708	1'979	1'235	287.4	***	***	***	
Canopy cover	29.4	79.7	53.6	12.2	16.4	87.3	61.4	12.3	***		***	
Ground vegetation cover	27.0	91.3	64.5	15.9	15.1	93.5	48.6	18.2	***	**	***	
Ground vegetation distribution	0.5	2.94	1.88	0.52	0.63	2.94	1.55	0.51	***		***	*
Ground vegetation height	8.4	55.8	30.1	10.7	7.6	52.8	26.1	10.7	*	*	*	*
Herb cover	3.4	54.2	23.0	12.1	3.4	50.5	17.9	10.5	**	**		
Fern cover	0	22.7	5.4	4.9	0	12.8	3.7	3.2	*		.	.
Bilberry cover	0	51.5	13.7	13.4	0	38.3	5.0	7.5	***	**	**	.
Basal branched trees	0	7.79	1.71	1.81	0	6.63	0.63	1.16	***	**	**	
Inner forest edges	0	9	1.16	1.68	0	6	0.64	1.01	*		*	
Total Rowans $\geq$ 3m	0	5.6	1.52	1.51	0	4.8	0.70	1.02	***	**	*	
Resource trees in tree layer	0	30.0	7.4	7.2	0	37.5	4.4	5.9	**	***		
Resource trees in shrub layer	0	92	26.0	20.1	0	58.3	14.3	12.4	***	**	**	
LiDAR variables - point cloud statistics												
Standard deviation 10th height percentile ($StdH_{10}$)	2.27	8.67	4.68	1.40	2.77	9.89	5.98	1.67	***	***	***	
Standard deviation 80th height percentile ($StdH_{80}$)	3.60	10.13	7.26	1.51	4.14	10.54	8.15	1.54	***	*	***	
Standard deviation 40th density percentile ($StdD_{40}$)	0.021	0.274	0.136	0.031	0.109	0.287	0.155	0.036	**	*	*	
Mean 20th density percentile ($MeanD_{20}$)	0.087	0.280	0.167	0.043	0.044	0.358	0.149	0.049	*	*		
Mean 70th density percentile ($MeanD_{70}$)	0.000	0.123	0.035	0.026	0.000	0.214	0.046	0.041	*	*		
LiDAR variables - FRAGSTATS												
Canopy height heterogeneity (CHH)	32'424	127'044	66'148	19'535	16'968	94'112	48'756	16'274	***	***	**	
Clumpyness of CHM class 4 (Clumpy_4)	0.486	0.895	0.764	0.089	0.543	0.933	0.828	0.065	***	***	**	
Shannon's diversity index for CHM classes 2 & 3 (ShDI2_3) neg. coif.!	0.095	0.693	0.521	0.147	0.236	0.693	0.572	0.109	*	*	.	
Shannon's diversity index for CHM classes 1 & 2 (ShDI1_2)	0.145	0.693	0.519	0.143	0.132	0.693	0.477	0.140	.			

B MATLAB code for processing LiDAR data

```
MATLAB code for processing Swisstopo LiDAR data
by F. Zellweger and F. Morsdorf, 15th Nov.  2011
fnames = {'all sampling unit ID's'};
fid = fopen('LiDAR_masterfile','w'); % open output file
fprintf(fid,'%s\n',['Header']); % define header of output file
numlay = 11; % number of bands/layers to be computed in als2multi()
for i = 1:  length(fnames) % for-loop for all sampling unit ID's
    DTM = load(['DTM_',fnames{i},'.xyz']); % load data files
    DOM = load(['DOM_',fnames{i},'.xyzt']);
    ptdensity = length(DOM)/1210000; % pts per m2 (1100m x 1100m)
    DOM = DOM(DOM(:,4)==0,1:3); % filter for veg.  points only
    [DTM] = raw2ras(DTM,1,2,'DTM'); % calculate DTM with res.  1m, moving window per.  2m
    DTM.z = inpaint_nans(DTM.z,4); % fill DTM holes with linear interpolation
    [DTM.X,DTM.Y] = meshgrid(DTM.x,DTM.y); % transform vectors x,y to arrays X,Y
      % terrain corrected vegetation height (vegh):
    vegh = DOM(:,3) - interp2(DTM.X,DTM.Y,DTM.z,DOM(:,1),DOM(:,2));
    CHM_p50 = [DOM(:,1),DOM(:,2),vegh(:,1)]; % vegh incl.  position, 50m buffer
    [CHM] = raw2ras(CHM_p50,4,8,'dsm'); % Canopy Height Model, res.  4m, mov.  wind.  per.  8m
    vegm = ones(size(CHM.z))*1; % classification of CHM, matrix
    vegm(CHM.z <=5) = 2; % classification of CHM, understory class
    vegm(CHM.z >5 & CHM.z <= 15) = 3; % classification of CHM, midstory class
    vegm(CHM.z >= 15) = 4; % classification of CHM, canopy class
    mat2asciigrid_vegm(['CHM4_',fnames{i}],vegm); % export CHM in ascii format
    MaxX = max(CHM_p50(:,1)); % get max and min of X and Y
    MaxY = max(CHM_p50(:,2));
    MinX = min(CHM_p50(:,1));
    MinY = min(CHM_p50(:,2));
      % cut the CHM data to 1 x 1 km:
    CHM_1km = CHM_p50(((CHM_p50(:,1) > (MinX+50)) & (CHM_p50(:,1) < (MaxX-50)) & ...
        ((CHM_p50(:,2) > (MinY+50)) & (CHM_p50(:,2) < (MaxY-50)))) ,1:3);
    vegh_2m = CHM_1km(CHM_1km(:,3)<2,1); % get all points below 2m
    penrate_2m = size(vegh_2m)/size(CHM_1km); % penetration rate at 2m
    raw.x = CHM_1km(:,1); % data structure for als2multi()
    raw.y = CHM_1km(:,2);
    raw.z = CHM_1km(:,3);
    [x,y,multih,multid,ras] = als2multi_fl(raw,10,numlay,50); % compute multi-band image
    for j = 1:numlay % inner for-loop to compute point cloud statistics
     minh(j) = min(min(multih(:,:,j))); % min value for each band/layer
     maxh(j) = max(max(multih(:,:,j))); % max value for each band/layer
     htemp = multih(:,:,j); % select multih for all 10x10m cells per band/layer
     htemp(isnan(htemp)) = []; % fill the nans with empty
     meanh(j) = mean(htemp); % calculate mean height percentiles per 1km2 per band/layer
     stdh(j) = std(htemp); % calculate std of height percentiles per 1km2 per band/layer
     mind(j) = min(min(multid(:,:,j))); % ditto for density related percentile values
     maxd(j) = max(max(multid(:,:,j)));
     dtemp = multid(:,:,j);
     dtemp(isnan(dtemp)) = [];
     meand(j) = mean(dtemp);
     stdd(j) = std(dtemp);
    end
    myformat = ['format type of output', e.g.  '%2.4f\n'];
    mydata = [str2num(fnames{i}),ptdensity,penrate_2m,stdh,stdd,meanh, etc.  % data body
    fprintf(fid,myformat,mydata); % print data to file
    save([fnames{i}]) % save MATLAB-file
end
fclose(fid) % close output file
```

C Personal declaration

I hereby declare that the submitted thesis is the result of my own, independent, work. All external sources are explicitly acknowldeged in the thesis.