MODELLING THE ABUNDANCE AND SPATIAL DISTRIBUTION OF MOUNTAIN GORILLA FORAGE SPECIES IN THE VIRUNGA MASSIF

PROVIDENCE AKAYEZU
February, 2016

SUPERVISORS:
Dr. I.C. van Duren
Dr. Ir. T.A. Groen

MODELLING THE ABUNDANCE AND SPATIAL DISTRIBUTION OF MOUNTAIN GORILLA FORAGE SPECIES IN THE VIRUNGA MASSIF

PROVIDENCE AKAYEZU
Enschede, The Netherlands, February, 2016

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Natural Resources Management

SUPERVISORS:
Dr. I.C. van Duren
Dr. Ir. T.A. Groen

THESIS ASSESSMENT BOARD:
Dr. Y.A. Hussin (Chair)
Dr. Ir. R.A. (Rolf) de By (External Examiner, ITC-Department of Geo-Information Processing)
Dr. I.C. van Duren (1st Supervisor)
Dr. Ir. T.A. Groen (2nd Supervisor)

ABSTRACT

Understanding the quantity and the spatial distribution of food species consumed by primate rare species is crucial in order to plan for conservation strategies. The population of 'critically endangered' mountain gorillas (*Gorilla beringei beringei*) has doubled since 1981 mainly due to success in conservation efforts in their natural habitat. This increase could intensify the competition for food and conflicts among gorilla groups. The mountain gorillas frequently visit areas with high abundance of food; which is mainly composed by herbs, vines and woody plants. Whilst the gorilla habitat suitability modelling under climate change scenarios alerts the future changes in gorilla habitat; no study has been done to show the current distribution of mountain gorilla most preferred food species. The latter being continuously monitored gives an idea on the movement of the mountain gorillas. This study was conducted in the Virunga massif and two datasets were used in the analysis. During October 2015, the researcher collected dataset 1 on gorilla food species abundance using the systematic sampling. The dataset 2 contained gorilla food species biomass and was collected using the random stratified sampling in 2009-2010 by Dr. Grueter Cyril in collaboration with the Dian Fossey Gorilla Fund International. On one hand, the two different datasets were used for testing the significant difference in each of the five most preferred gorilla food species abundance or biomass between vegetation types using the ANOVA. Similarly, the relationship between gorilla food species abundance or biomass and environmental conditions was examined through the multiple linear stepwise regression. On the other hand, the field measured canopy was correlated to the Aster band reflectance values to produce the Virunga forest canopy cover which is one of the predictors of the plant species occurrences. Finally, both datasets were combined and abundances/biomass converted into presence/absence which together with a set of seven environmental variables formed inputs to the 'BRT' species distribution model. When it was significant, the ANOVA test showed differences in gorilla food biomass between vegetation types in the Virunga massif. The higher abundance (> 40%) or biomass (> 5 g/m^2) of both *Rubus* spp. and *Galium* spp. were found in the Hagenia-Hypericum, the brush ridge and the sub-alpine vegetation zones; and the two species could be found in the alpine zone. The *P. linderi*, the *C. nyassanus* and the *L. alatipes* higher abundance (> 40%) or biomass (> 5 g/m^2) were observed in the herbaceous, the Hagenia-Hypericum and the brush ridge. These three species were totally absent in either the sub-alpine or alpine vegetation zones. Although significant, the Aster band reflectance values poorly predicted the variation in Virunga forest canopy cover (R^2=0.128; $p<$ 0.05). The significant stepwise regression model ($p<$ 0.05) showed the increase in both *Galium* spp. and *Rubus* spp. abundance or biomass in higher elevations whereas the *P. linderi*, the *C. nyassanus* and *L. alatipes* biomass declined in higher elevations. Likewise the BRT model with accuracies greater than random guess (AUC > 0.55) predicted higher probabilities (> 0.60) of occurrence of *Galium* spp. and *Rubus* spp. even on the volcanoes peak in the Virunga massif. In contrast, the *P. linderi*, *C. nyassanus* and *L. alatipes* had very low probabilities of occurrence (< 0.20) in higher elevations. The elevation and eastness were the two most important predictors of the five species occurrence with a common trend of higher probabilities of occurrence on the western-facing slopes. The findings of this study suggest the follow up on the changes in gorilla suitable areas especially the frequency of their visits in the higher elevations which were revealed by previous studies to be unsuitable for gorillas.

Keywords: Mountain gorilla, five preferred food species, abundance/biomass, Multiple Linear Regression, Boosted Regression Tree model, Virunga massif

ACKNOWLEDGEMENTS

"*Je te louerai l'Eternel de tout mon coeur, je raconterai toutes tes merveilles*". Psaumes 9

First, I would like to express my deep gratitude to the Dutch Government through the Netherlands Fellowship Program (NFP) for granting me the scholarship to boost my conservation biology career.

Second, I would like to address special thanks to my thesis supervisors Dr. Iris van Duren and Dr. Ir. Thomas Groen for their tireless support and guidance that helped me throughout the whole process of the scientific research just from the beginning until the final compilation of this MSc. thesis. Thank you so much for always opening your doors for me in case I ran into trouble or panic about my research.

Sincere thanks to the NRM department staff at ITC; the course Director Mr. Raymond Nijmeijer; you helped me right when I needed help most. Thank you Dr. Wieteke Louise for connecting me to your close friend in IGCP and discussion about my research topic.

The fulfilment of this study could have been a fantasy without the contribution from individual awesome people and institutions. Thanks to the Karisoke Research Centre, for inviting me in a seminar through which I met gorilla scientists. Dr. Martha Robbins and Dr. Cyril Grueter; you have been great to me when you accepted to share with me your data to be used for my thesis. Thanks Dr. Winnie Eckardt for the short discussion we had before the seminar when you told me 'today you're lucky'.

With pleasure I thank the Rwanda Development Board; conservation division personnel: Mr. Télesphore Ngoga, Dr. Tony Mudakikwa, Mr. Abel Musana, Mr. Prosper Uwingeli who were always in touch with me, offered me the research permission and made all my fieldwork arrangements. Thanks to RDB rangers and trackers who took me in the volcano ravines and summits, crawling together in a dense rainforest.

I would also like to acknowledge Mr. Jean Chrysostome Sehene and Mr. Ernest Bucyayungura for advising me to apply for this scholarship and recommending me. I always recognize your encouragement, friendship and tell you how much they have meant to me.

Special appreciation to my Lovely NRM/GEM classmates; 'you guys rock'. 18 months we spent together, I am very proud to have been part of a wonderful group of people I admire. Four of us were Rwandans; thanks Marceilline, Aline and Maurice. From the bottom of my heart, I thank Tasi, Agnes, Sadadi, Aristotle, Phanintra and Laura; you have been extremely supportive through difficult moments.

Many thanks to the Rwandan colleagues at ITC, residents in Enschede; you made me feeling at home while staying in the Netherlands. Thank you Nyandwi Elias for spending one night reading and commenting on my research proposal.

Finally, cordial thanks go to my family; my parents, my brother Bunani, my sister Adrie, my brother-in-law Lucien; your endless prayers and love have made me a strong person and I will forever be grateful.

Providence Akayezu
Enschede, the Netherlands
February, 2016

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ACRONYMS

ANOVA	Analysis of Variance
ASTER	Advanced Spaceborne Thermal and Reflection Radiometer
AUC	Area Under the ROC Curve
BRT	Boosted Regression Tree
CART	Classification And Regression Trees
DBH	Diameter at the Breast Height
DEM	Digital Elevation Model
DFGFI	Dian Fossey Gorilla Fund International
DN	Digital Number
DRC	Democratic Republic of Congo
ECW	Enhanced Compression Wavelet
ETM+	Enhanced Thematic Mapper Plus
EVI2	Two bands Enhanced Vegetation Index
GAM	Generalized Additive Models
GIS	Geographic Information System
GLM	Generalized Linear Models
GPS	Global Positioning System
IGCP	International Gorilla Conservation Programme
IPAQ	Pocket PC
ITC	Faculty of Geo-Information Science and Earth Observation
IUCN	International Union for Conservation of Nature
KRC	Karisoke Research Centre
MaXent	Maximum Entropy
MINECOFIN	Ministry of Finance and Economic Planning, Rwanda
MLR	Multiple Linear Regression
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
RDB	Rwanda Development Board
REMA	Rwanda Environment Management Authority
RNRA	Rwanda Natural Resources Authority
ROC	Receiver Operating Characteristic
RS	Remote Sensing
SDM	Species Distribution Modelling
SWIR	Short Wave Infrared
TIR	Thermal Infrared
TOA	Top of Atmosphere
TSS	True Skills Statistics
VI	Vegetation Indices
VIF	Variance Inflation Factor
VNIR	Visible-Near Infrared
VNP	Volcanoes National Park

1. INTRODUCTION

In this chapter, the background, relevance and conceptual diagram of the study are described. The chapter concludes with the research objectives, research questions, hypotheses and assumptions.

1.1. Background

The mountain gorilla (*Gorilla beringei beringei*) are "critically endangered" primate species whose habitat ranges up to 4507 m higher than any other gorilla species (Spinage, 1972). The Virunga massif shared between Rwanda, Uganda and Democratic Republic of Congo (DRC) is one of the suitable habitat for mountain gorillas. The latter are mainly folivorous because their habitat lacks fruits. Two aspects make the Virunga mountain gorillas being easily followed and observed: when they feed and move in the dense undergrowth of montane rainforest; a well-marked path of trampled vegetation is left behind, but also the rugged terrain allows visibility while maintaining a distance between the observer and the gorillas (Robbins et al., 2005).

In the early 1970s, the mountain gorilla census showed a decline in their population from 450 to about 275 animals in the Virunga. The main cause was habitat conversion into settlement and cattle grazing which occurred between the 1958 and 1973, especially on the side of Rwanda (Spinage, 1972). At the end of 1980s, conservation efforts led to the increase of the Virunga mountain gorilla population from 250 to 480 gorillas counted in 2010. Although political unrest and violence characterized the East African region from 1991 to 1998; the research and conservation activities have rebuilt completely and the current mountain gorilla population continue to increase (Gray et al., 2013). However, the gorilla population dynamics may affect directly their food abundance (Grueter et al., 2013).

Studies have been done on the spatial and seasonal variation in gorilla food species abundance or biomass in the Virunga protected area (Watts, 1984; Watts, 1998c; Watts, 1998a; Watts, 1998b; Vedder, 1984; Plumptre, 1991; Grueter et al., 2013). On one hand, using the stratified-random sampling, food abundance measurement techniques included quantifying the number of emergent stems, measure the leaf length, stem length and cover rating (1-5). The abundance was then converted into biomass by first harvesting the plant stem, leaves and use regressions to transform the leaf length into leaf mass, etc. All the plants parts harvested were sundried until there is no further loss in weight and then weight again to get the gorilla food dry biomass in grams per square meter. On the other hand, the food species frequency was calculated by summing the number of plots in which a particular food is occurring. Most of these studies were conducted in the Karisoke study area located in the saddles between the Karisimbi, Mikeno and Bisoke Volcanoes. Food species abundance in space was mainly estimated along the altitudinal gradient and vegetation zones (Grueter et al., 2013).

The abundance and distribution of food influence the mountain gorilla ranging patterns and a high frequency of visits of areas with high food density was noticed (Vedder, 1984). The changes in food resources availability together with a reduced ecosystem carrying capacity may result from increased animal growth rate.

For example Grueter et al. (2013) found that two of the most important gorilla food species decreased while three increased in availability i.e. the number of percentage plots it was observed from 1989 to 2010. In addition, altitudinal changes in gorilla food density distribution were observed (Grueter et al., 2013).

Taking into consideration the mountain gorilla population growth along with the vulnerability of their habitat to anthropogenic pressure and climate change (Foster, 2001); it is crucial to continuously monitor the abundance and the spatial distribution of gorilla food species (Grueter et al., 2013).

Given the advancement of new powerful statistical techniques together with Geographical Information System (GIS) and Remote Sensing (RS) tools; it is possible to relate the species occurrence data at known locations with ecological characteristics of those locations. Eventually, the probability of occurrence maps as well as the best predictors of the species occurrence are produced (Elith & Leathwick, 2009; Duque-Lazo et al., 2016; Guisan & Zimmermann, 2000).

1.2. Relevance of the study

The mountain gorilla eat over 200 plant species but about 20 species make up the majority of their diet. Herb foods containing high protein and little fibre content are preferred by mountain gorillas; and they are distributed throughout their habitat (Ganas et al., 2008). The improved positive human awareness, veterinary care and continuous close monitoring of wildlife has led to the increase in number of habituated mountain gorillas in the Virunga protected area from ca 250 to 400 individuals since 1981 (Robbins et al., 2011). This mountain gorilla population growth could lead to competition for food among herbivores in the Virunga massif. Moreover, climate change is supposed to affect the East African Tropical rainforests including the mountain gorilla habitat (Foster, 2001). Under global warming conditions, species normally shift to higher latitudes and elevations or migrate within the extent of their current range. Indirect effects include changes in the amount and distribution of available food items or changes that provide favourable conditions for invasive species (Belfiore et al., 2015). For example shifts towards higher altitudes of one the dominant vegetation type (*Hagenia abyssinica*) in the Volcanoes National Park were observed. The opposite situation happened for one of the most preferred gorilla food (Grueter et al., 2013). Although mountain gorillas are flexible to adapt their diet; they rarely try new plant species to eat; which means that they are tightly tied to the location of their primary food plants. The modelling of the current and future mountain gorilla suitable habitat under climate scenarios has already been carried out (Belfiore et al., 2015; van Gils & Kayijamahe, 2010). Both studies used gorilla presence only data together with environmental variables input to the MaXent algorithm; and the models performed with adequate accuracies (Area Under the ROC Curve > 0.70). Nevertheless, the current spatial distribution of the most preferred mountain gorilla food species is unknown. The emphasis of the present research is first to prepare both GIS and RS variables including the slope aspects parameters such as the eastness and northness as well as the canopy cover and vegetation types; and link them to the gorilla food abundance or biomass. Second, these variables together with species presence/absence data constitute the inputs to the "Boosted Regression Tree" in order to produce the spatial distribution maps of five most preferred food species by mountain gorillas. Those are *Galium* spp., *Carduus nyassanus*, *Peucedanum linderi*, *Laportea alatipes* and *Rubus* spp.. The results are useful to compare the mountain gorilla current home ranges with the distribution of their food. In addition for upcoming researches, the knowledge on the changes in the distribution of gorilla food species is a good indicator of the future suitable mountain gorilla habitat.

1.3. Mountain gorilla food species and relationships with their environment: Conceptual diagram

Species distribution modelling consists of relating the species occurrence, abundance or absence data at known locations with spatial or ecological characteristics of those locations (Elith & Leathwick, 2009; Duque-Lazo et al., 2016). The earth biotic and abiotic factors define the species' geographical range. For instance elevation, vegetation cover, substrate (e.g. soil) were found to be the best predictors of plant taxa distribution in the Majella National Park; Central Italy (van Gils et al., 2012).

In the Virunga massif, the gorilla food species abundance or presence can have a straight relationship with environmental variables; and the latter are either directly or indirectly linked to remote sensing or GIS. In the current study, four categories were distinguished: the topographic, climate, forest structure or substrate and variables extracted from the satellite imagery. The topography group includes slope, aspect, elevation and solar radiation; all obtained from a Digital Elevation Model and prepared in a GIS software. Temperature and rainfall form the climate group; while vegetation types, forest canopy cover, tree height, stem density and tree diameter at the breast height constitute the vegetation structure division. This third group of variables are collected from field while the band reflectance values, the NDVI and the EVI2 can be extracted from the satellite imagery.

Elevation is an important parameter determining vegetation zones in the Virunga massif. The topography, soil conditions and particularly altitude create variation in local vegetation composition (Watts, 1984). Using the MaxEnt model; vegetation types, solar radiation and slope were found to be crucial predictors of mountain gorilla habitat suitability in the Virunga massif (Kayijamahe, 2008). Altitude (elevation) was also used by Grueter et al. (2013) to predict the gorilla food availability in the Volcanoes National Park, Rwanda. However, the collinearity was observed between elevation and climate variables: minimum or maximum temperature, average temperature and precipitation (Kayijamahe, 2008). Thus only elevation can be used as a proxy of climate variables but also climate data from the WorldClim are too coarse to be used with the 15 m resolution Aster imagery and a 454 km^2 study area.

The vegetation index is a good indicator of a healthy or stressed vegetation (Fatiha et al., 2013). The NDVI extracted from the Aster DN values was used as one of the environmental layers to model the plant distribution in Majella National Park (van Gils et al., 2012). The giant panda forage abundance in China was predicted using both the NDVI and EVI vegetation indices (Wang, 2009). Hence, the two vegetation indices can be used as predictors of mountain gorilla food species biomass or abundance.

The plants growth depends on the amount of light reaching their vegetative parts (leaves, stems); this light varies with the vegetation structure. For understory plant species, Martens et al. (2000) found that the mean understory light decreased with increased overstory cover and sensitive to changes in canopy height. The dominant canopy plant species determine the spectral signature and thus can be mapped using direct remote sensing approaches (Joshi et al., 2006). The understory species, however do not show any spectral signature and their mapping can be done based on the knowledge of their ecological relationship with the environment (Joshi et al., 2004). Multiple linear regression was used to predict the forest canopy density predicted by the Landsat ETM+ band reflectance values (Joshi et al., 2006). Moreover, Aster band reflectance values as well as the canopy height and cover were used to predict the presence and abundance of vascular epiphytes in Nyungwe National Park, Rwanda (Nyandwi, 2008). Therefore, a satellite imagery band reflectance values can be used to predict the tree canopy cover in the Virunga massif and the canopy can be input for the gorilla food species presence modelling.

The topographic factors affect mountain forests through their direct influence on light intensity, soil and atmospheric moisture, soil and air temperature and wind velocity (Måren et al., 2015; Austin, 2002). The slope and aspects were found to modify the vegetation composition and dynamics in motorway slopes of Valencia, Spain (Bochet & García-Fayos, 2004). Additionally, Piedallu & Gégout (2008) suggested using the solar radiation combined with slope-aspect transformations for plant species distribution at a local scale. Generally in the southern hemisphere, maximum temperatures and clear mornings followed by cloudy afternoons make the eastern-facing slopes being drier while the western slopes receive early morning clouds which make them moist. The south-facing are then moist in contrast to the north-facing slopes and the opposite phenomena happens in the northern hemisphere (Smith, 1977; Måren et al., 2015). Thus, the topography variables can be used for predicting both the gorilla food species abundance or biomass and spatial distribution.

Forest structure parameters such as tree diameter, stem density and height may be related to the abundance of the understory plant species including lianas. The presence of an understory graminoid species was positively related to the mean diameter and number of stems of *Pinus ponderosa* in Arizona (Naumburg & Dewald, 1999). Accordingly, the tree diameter, height and stem density can be related to the abundance or biomass of mountain gorilla food species.

Relationships between the climate, topography and vegetation structure in the Virunga massif are shown on Figure 1 below.

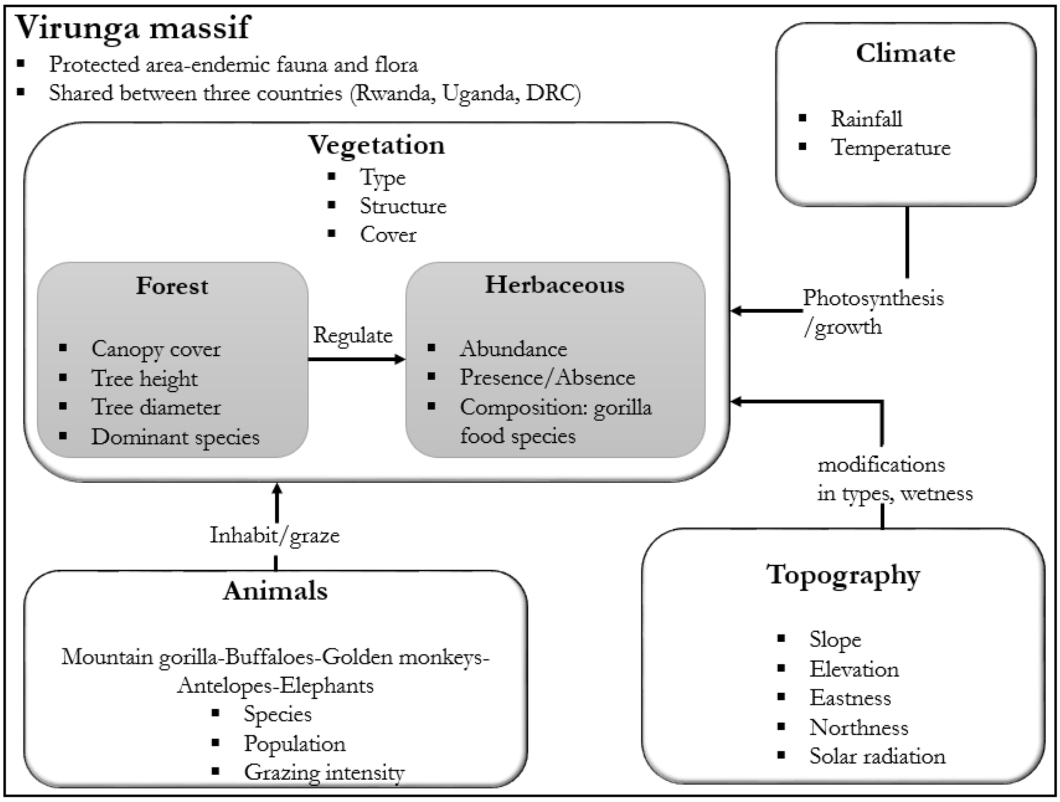


Figure 1: Conceptual diagram showing interactions between the vegetation, animals and topography together with climate variables.

1.4. Research objectives

1.4.1. General objective

The overall objective of this study is "to map and model the abundance and spatial distribution of five (except bamboo) most preferred gorilla forage species in relation to biophysical variables". The Table 1 below depicts specific objectives and corresponding research questions.

1.4.2. Specific objectives and research questions

Table 1: Research objectives and questions

Research objectives	**Research questions**
1. To examine the significant differences in abundance or biomass of each of the five gorilla food species between vegetation types.	How do gorilla food species abundance or biomass vary within each vegetation type?
2. To assess the relationship between the forest canopy cover, dbh, tree height, stem density, altitude, eastness, northness, solar radiation, the slope and the abundance or biomass of gorilla forage species.	How strong is the relationship between individual gorilla food species abundance or biomass and the forest canopy cover, dbh, stem density, tree height, altitude, slope, northness, eastness and solar radiation?
3. To determine the relationship between the forest canopy cover and the nine Aster imagery band reflectance values.	What is the relationship between forest canopy cover and nine Aster bands reflectance values?
4. To examine the relationship between the gorilla forage species abundance or biomass and the NDVI, EVI2	What is the relationship between the Aster NDVI, EVI2 and individual gorilla food species abundance or biomass?
5. To model and map the distribution of each individual mountain gorilla food species.	a. How accurate can the Boosted Regression Tree discriminate between the presence/absence of each of the five gorilla food species? b. What are the most important environmental variables predicting the occurrence of each of the five gorilla food species?

1.5. Research hypotheses

Hypothesis 1

H_0: There is no significant difference in individual gorilla food species abundance or biomass between vegetation types.
H_1: There is a significant difference in individual gorilla food species abundance or biomass between vegetation types.

Hypothesis 2

H_0: There is no significant relationship between the canopy cover and the Aster imagery band reflectance values.
H_1: There is significant positive relationship between the canopy cover and the Aster imagery band reflectance values.

Hypothesis 3

H_0: Each of the five gorilla food species abundance or biomass declines in higher elevations, western facing-slopes or small tree densities.
H_1: Each of the five gorilla food species abundance or biomass increases in higher elevations, western facing-slopes or small tree densities.

Hypothesis 4

H_0: The abundance or biomass of individual gorilla food species is poorly explained by NDVI than EVI2.
H_1: The abundance or biomass of individual gorilla food species is better explained by NDVI than EVI2.

Hypothesis 5

H_0: The Boosted Regression Tree model cannot discriminate each of the gorilla food species presence/absence with an accuracy greater than random guess.
H_1: The Boosted Regression Tree can discriminate the presence/absence of each of the gorilla food species with an accuracy greater than random guess.

1.6. Assumptions

The two Aster satellite images used in this study for vegetation mapping were acquired February 2005 and June 2006 respectively. These images were selected because of the higher resolution (15 m) and few or absent clouds or haze. The 2006 Aster image was not covering the eastern part of the study area while the 2005 was covering the whole study area but had some clouds; thus vegetation classification was done for each image separately and then combined them to make a final vegetation map. For extracting the band reflectance values and calculating vegetation indices, the 2006 Aster image was then used alone as it had very few clouds with only 3 years time difference with the Dian Fossey Gorilla Fund International (DFGI) data. It was assumed that no major changes occurred in the Virunga massif area since 2006. Therefore, variables extracted from these images are reliable to be linked with 2009-2010 field data collected by the DFGFI, but also data collected September-October 2015 by the researcher. Moreover, more than 80% of the field points were collected on the side of Rwanda. Given that in the Virunga massif the altitude defines environmental conditions; it was assumed that land cover types are similar on both sides of the volcanoes and thus it was possible to make analysis including the Democratic Republic of Congo and Ugandan side of the park.

2. LITERATURE REVIEW

In this chapter, the key literature related to vegetation structure and mountain gorilla habitat, feeding ecology, remote sensing for habitat mapping and the methods used to map the spatial distribution of the gorilla forage species are emphasized.

2.1. Vegetation structure and mountain gorilla habitat suitability

The Bwindi Impenetrable National Park and the Virunga massif are the two remnant forests suitable for the mountain gorilla habitat (Figure 3). The Virunga massif is a chain of six dormant volcanoes which covers an area of approximately 450 km^2. Being a montane rainforest; variation in topography, soil conditions and particularly altitude are limiting factors for vegetation structure and composition in the Virunga massif (Weber & Vedder, 1983). Table 2 summarizes the altitudinal zones and corresponding vegetation types characteristics in the Virunga protected area.

The mountain gorilla habitat use depends on variation in food distribution (Watts, 1998b). In the Virunga massif, the mountain gorilla forage species are unevenly distributed in various vegetation types (Watts, 1998a). By considering the quantity and quality of food availability; four major vegetation belts are important for mountain gorillas. Those are the bamboo zone (2,500 m-2, 800 m), *Hagenia-Hypericum* (2,800 m-3,200 m), herbaceous (2,800 – 3,300 m) and brush-ridge (3,000 – 3,300 m). Dominantly occurring in the west of the park, the *Hagenia- Hypericum* vegetation includes trees with open canopies; so that herbaceous species proliferate on the forest floor. The bamboo zone forms completely closed canopies and thus support little understorey growth. The mountain gorillas seasonally depend on the bamboo vegetation; because they feed on bamboo shoots which are available in this zone in October, November and early December of each year (Vedder, 1984). The brush-ridge and herbaceous vegetation are dominated by tall herbs and scattered shrubs or trees (Weber & Vedder, 1983; Fossey, 1974; Plumptre, 1991; Grueter et al., 2013). Although *Hagenia abyssinica* and *Hypericum revolutum* forests are overlapping in the Virunga massif; in some areas they can be distinguished from each other on the basis of dominance between the two species. The *Hypericum* shrubs dominate on the moderate to higher elevations where dense closed canopies are absent; and they support a variety of vines which are eaten by the gorillas. The *Hagenia abyssinica* giant trees with closed canopies are found on the Karisimbi, the Mikeno and the saddle area in between the two volcanoes together with Bisoke (Dondeyne et al., 1993).

There is a correlation between mountain gorilla suitable areas and food abundance (Vedder, 1984). The mid-altitudes (2, 500 – 3,500 m) were found to be suitable (Figure 3) for mountain gorillas (van Gils & Kayijamahe, 2010; Belfiore et al., 2015) and this finding was in accordance with gorilla home ranges recorded in 1976 – 1978 (Figure 2). The mountain gorillas avoid higher altitudes because of the very cold and wet conditions which can be sources of respiratory diseases while the lower altitudes are close to human impacts (Watts, 1998a; Kayijamahe, 2008).

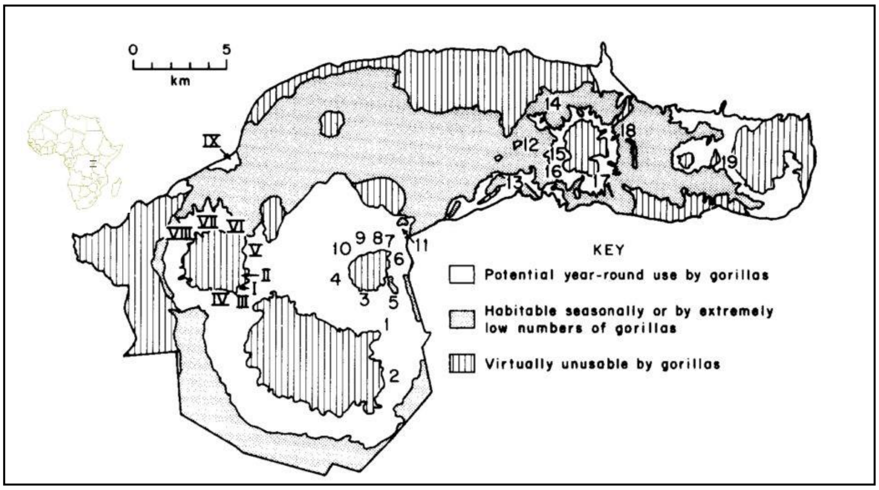


Figure 2: The suitability of different areas to harbour mountain gorillas in the Virunga massif[a].

[a]The numbers indicate different gorilla groups recorded in the 1976 – 1978 census. Map adopted from Weber & Vedder (1983).

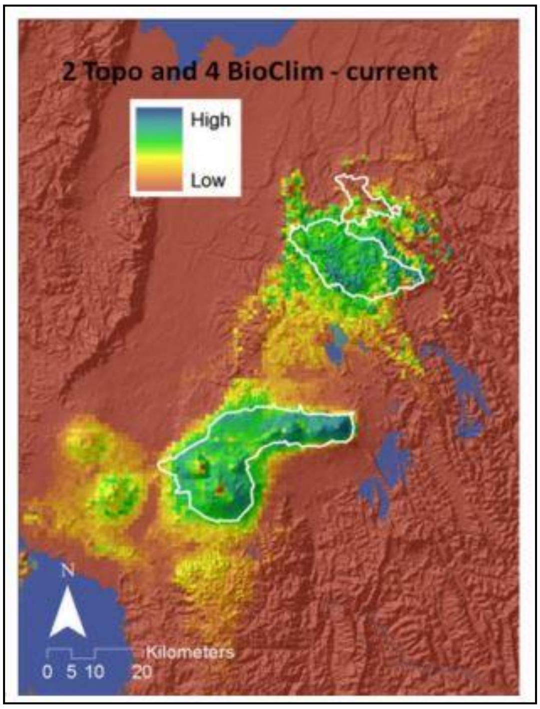


Figure 3: Bwindi (on top) and Virunga (down) current mountain gorilla range predicted by MaXent using 2 topographic and four BioClim variables[a].

[a]Map taken from Belfiore et al. (2015).

Table 2: vegetation zones in the Virunga region

Vegetation zone	Characteristics	Land cover types description	Altitude (m)
Alpine	Grasses, mosses and lichens, *Dendrocenescio*, giant *Lobelia,* it occupies 6 per cent of the park.	• Sparse vegetation area including transitional stages from sub-alpine zones to areas where most forms of plant life are extremely limited	Above 3,600
Sub-alpine	*Philippia johnstonii*, *Erica arborea*, Giant *Lobelia* Giant *Scenecio*	• Usnea lichens present • grasses	3,200-3,600
Brush ridge	• Occurs on volcano slopes, • forms the edges of deep ravines	• Certain shrubs (favourite mountain gorilla food). • The main shrub species are *Rubus kirungensis* and *Rubus runssorensis*	3,000 – 3,300
Herbaceous	Consists of dense tall herbs with no tree cover	Favourite mountain gorilla food	2,800 – 3,300
Hagenia-Hypericum forest	*Hypericum revolutum*, *Hypericum absi*, *Hagenia abyssinica*	• *Hagenia abyssinica* trees (height between 15 to 24 m) and • *Hypericum revolutum* (saplings to full-grown tree with a height equal to 15 m).	2,800-3,200
Bamboo	*Arundinari alpina*	• Bamboo strips • Closed canopy • Covers 35 % of the park	2,500-2,800
Disturbed woodland	Areas of regenerating forest that were cultivated in Mgahinga	• Young *Neoboutonia* trees growing • Very little or absent forest undergrowth	2,300-2,800
Mixed forest	Moist semi-deciduous forests with broad leaves	• *Dombeya goetzenii* (moist montane forest with *Mimulopsis* shrub) • Tall trees (height > 20 m) with broad leaves • Large part cleared for agriculture • Occupies 20 % of the park	1,600-2,500
Grassland	Areas dominated by grass	Meadow and savannah	Occurs at various altitudes
Swamp	Marshy or boggy areas	In the saddles between volcanoes	Occurs at various altitudes

2.2. Mountain Gorilla feeding ecology

Mountain gorillas are primarily herbivorous; their diet is largely made up of leaves, shoots, pith of terrestrial herbs and fruits; rich in both sugar and protein (Ganas et al., 2008). All these gorilla food items are abundant and widely distributed through their habitat (Doran & Mcneilage, 1998). Mountain gorillas don't drink water from free-standing water sources, but rather obtain water from food sources (Advani, 2014). Fossey (1974) and Watts (1984) made a list of food resources commonly consumed by mountain gorillas. They are described in the Table 3 below:

Table 3: Some food resources eaten by mountain gorillas

Plant species		
Species name	Family	Part eaten[a]
• Nettles: *Laportea alatipes*	Urticaceae	Lv, St, Rt
• *Galium* spp.	Rubiaceae	Whole
• *Hagenia abyssinica*	Rosaceae	Dead wood, Bk, Pi
• Bamboo: *Yushania alpina*	Poaceae	Shoots
• *Lobelia giberroa*	Lobeliaceae	Rt, Pi, Bk
• *Senecio trichopterygius*	Asteraceae	St
• *Carex bequaerti*	Cyperaceae	Lv, Infl
• *Cynoglosum lanceolatum*	Boraginaceae	Rt, St
• *Droquetia iners*	Urticaceae	Lv, St
• *Rumex ruwenzoriense*	Polygonaceae	St
• Thistle: *Carduus nyassanus* *Carduus leptocanthus* *Carduus kikuyorum*	Asteraceae	Lv, Fl, St, Rt
• Celery: *Peucedanum linderi* *Peucedanum kertstenii*	Apiaceae	St, Rt
• *Hypericum revolutum*	Hypericaceae	Dead wood, Bk, Rt, Lv
• *Urtica massaica*	Urticaceae	Lv, St, Rt
• Blackberries: *Rubus runssorensis* *Rubus kirungensis*	Rosaceae	Lv, Fr, St
• *Vernonia adolfi-frederici*	Asteraceae	Fr, Pi, Rt, Fl
Special foods		
• Ants • Dung • Milk • Mushroom • Placenta • Soil • Water		

[a]Lv: Leaves; St: Stem; Rt: Roots; Bk: Bark; Pi: Pith; Infl: Inflorescence; Fl: Flower; Fr: Fruit

To quantify the most relevant gorilla food species, the dietary importance of the food species has to be determined.

The dietary importance refers to the time spent by gorillas foraging on a particular plant (Figure 4); it is calculated from instantaneous activity samples taken at 10-min intervals during focal follows.

Figure 4: A mountain gorilla feeding on "*Carduus nyassanus*". Photo taken during fieldwork in Volcanoes National Park, Rwanda.

2.3. Five most preferred mountain gorilla food species

Shown on Figure 5, the six most preferred gorilla forage species are: *Galium* spp. (vine), *Carduus nyassanus* (tall herb), *Peucedanum linderi* (tall herb), *Yushania alpina* (bamboo shrub), *Laportea alatipes* (tall herb) and *Rubus* spp. (shrub). These plant species make up more than 70% of the mountain Gorilla diet (Watts, 1984; Grueter et al., 2013). Their ecology is described in Table 4 below.

Table 4: Some ecological aspects of five most preferred (except bamboo) gorilla food species.

Plant	Ecology (habitat)	Reference
Rubus spp.	• High affinity for highly disturbed areas • Establishing early and dominates sites after disturbance • Dominate forest gaps and competitively inhibit the recruitment of other species • Altitude 3300 m	(Kern et al., 2013)
Galium spp.	• Can survive in canopy shading environment • Occurs on rocky open habitats • Drought resistant	(Tang et al., 2014); (Raevel et al., 2013)
Peucedanum linderi	• Well adapted to grow in low light conditions • Occurs by mountain streams	(Garrison et al., 2005)
Laportea alatipes	• Occurs in undergrowth, usually near water	
Carduus nyassanus	• Swampy or moist grassland/moorland • Often in moist sites or bogs • Bamboo margins • Altitude: 1650 m- 3150 m	(Luke, 2010)


Galium spp. *Rubus spp.* (blackberry)


Carduus nyassanus (thistle) *Laportea alatipes* (nettles)


Yushania alpina (bamboo) *Peucedanum linderi* (celery)

Figure 5: Six (including bamboo) most preferred food species by mountain gorillas.

Photo taken during October 2015 fieldwork in Volcanoes National Park, Rwanda.

2.4. Habitat description from satellite imagery

Satellite sensors have the potential to measure the features on the earth's surface and enable detecting and quantifying changes on the earth's environment. The Advanced Spaceborne Thermal and Reflection Radiometer (ASTER) on-board Terra was launched the 18th December 1999 by the NASA in collaboration with the Japan Ministry of International Trade and Industry. It is a representative of multi-spectral and high to medium resolution stereo imagery from space (Table 5). The VNIR bands are very useful to acquire information on vegetation (Yamaguchi et al., 1998). The Aster data also provides a Digital Elevation Model (DEM) which is used to extract topography variables.

Table 5: Aster bands description

Subsystem[a]	Band No.	Spectral range (μm)	Spatial Resolution (m)
VNIR	1	0.52-0.60	15
	2	0.63-0.69	
	3N (Nadir)	0.78-0.86	
	3B (Backward)	0.78-0.86	
SWIR	4	1.60-1.70	30
	5	2.145-2.185	
	6	2.185-2.225	
	7	2.235-2.285	
	8	2.295-2.365	
	9	2.360-2.430	
TIR	10	8.125-8.475	90
	11	8.475-8.8.825	
	12	8.925-9.275	
	13	10.25-10.95	
	14	10.95-11.65	

[a]VNIR stands for 'Visible Near Infrared '; SWIR: 'Short-Wave Infrared'; TIR: 'Thermal Infrared' bands.

The digital image classification is a useful method in generating land cover data using either supervised or unsupervised classification of satellite data. The process is based on the different spectral characteristics of different features observed on the earth surface. The supervised classification is powerful because the operator defines the spectral characteristics of the classes by identifying sample areas (Richards & Jia, 2006). In addition, the operator should be familiar with the area of interest or carry out a fieldwork to relate what is observed on the image and the ground truth (Rogan, 2004).

Remotely sensed spectral vegetation indices (VIs) are spectral transformations of two or more bands designed to enhance the contribution of vegetation properties and allow reliable spatial and temporal inter-comparisons of terrestrial photosynthetic activity and canopy structural variations. They are used to monitor seasonal, inter-annual, and long-term variations of vegetation structural, phenological, and biophysical parameters (Huete et al., 2002). The Normalized Difference Vegetation Index (NDVI) is one of the commonly used as the best indicative factor of plant growth state and vegetation coverage (Goward et al., 1991). However, the NDVI is sensitive to soil background brightness and it tends to saturate in dense forest (Bausch, 1993; Huete et al., 1985); therefore a two bands Enhanced Vegetation Index (EVI2) can be employed in addition to NDVI.

2.5. Species distribution modelling: the Boosted Regression Tree model

The Boosted Regression Tree (BRT) is a machine learning algorithm that improves the accuracy of a single tree model through fitting several models and combining them for prediction. The BRT model consists of two components: regression trees and boosting (De' Ath & Fabricius, 2013).

2.5.1. Decision trees

The Classification and Regression Trees (CART) decision tree is a binary recursive partitioning procedure which can process both continuous and nominal data. The beginning consists of the root node, the data are split into two children, and each of the children is in turn split into grandchildren (Figure 6). The split is selected based on the reduction of the error of the tree of two resulting groups. Trees are grown until no further splits are possible and the maximal-sized tree is then pruned back to the root. The CART mechanism produces a sequence of trees each being a candidate to be the optimal tree. Cross-validation is used to obtain a 'honest tree size' or best estimated predictive single tree which has the smallest estimated error (De'Ath and Fabricus, 2000). With regression trees both categorical and continuous variables can be handled, the CART algorithm identifies the significant variables and at the same time eliminates non-significant ones; but also isolate outliers in a separate node. Moreover, the CART results do not change even if one or several independent variables undergo logarithm or square root transformations (De'ath, 2007). Nevertheless, modifications in the training sample could lead to the increased tree complexity or in splitting variables. In addition, the CART splits only by one variable; which result in poor handling of data with complex structure. To solve the problems resulting from regression trees, a 'boosting' approach is used (De'ath, 2007).

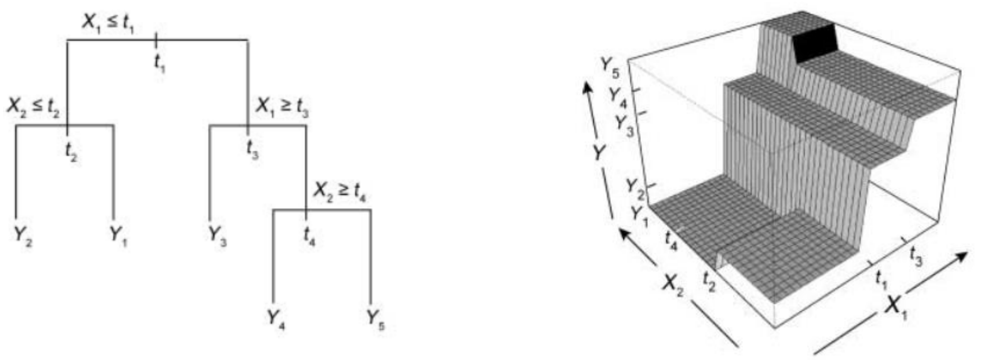


Figure 6: A single decision tree (left) with a response *Y*, two predictor variables, X_1 and X_2, and splits points t_1, t_2, etc. and its prediction surface (right)[a].

[a]Figures from Hastie et al. (2001).

2.5.2. Boosting

The boosting is a forward, stagewise procedure in which the models (decision trees) are fitted iteratively on the training data and the appropriate method is used to increase the accuracy of existing collection of trees. Boosting is a numerical optimization technique for minimizing the loss function by adding, at each step, a new tree that best reduces the loss function (i.e. a measure such as deviance which represents the loss in predictive performance).

The process is stagewise; which means that the existing trees are left unchanged as the model is enlarged; only the fitted value for each observation is re-estimated at each step to reflect the newly added tree. The final BRT comprises of hundreds to thousands trees that constitute a regression model where each term is a tree. The BRT is stochastic as it includes a random or probabilistic component. This means that even if a random seed initially set, the final models are subtly different each time they are run.

The tree complexity (*tc*) which controls whether interactions are fitted and the learning rate (*lr*) which determines the contribution of each tree to the growing model; are two important parameters which determine the number of trees required for optimal prediction. The fitted values in the final model are computed as the sum of all trees multiplied by the learning rate and they are more accurate than those from a single decision tree (Elith et al., 2008).

3. MATERIAL AND METHODS

3.1. Overview

The purpose of the current study was to model the mountain gorilla food species abundance or biomass and spatial distribution in the Virunga protected area. On one hand, the statistical significance and strength of the relationship between the gorilla food species abundance or biomass and biophysical factors was assessed. The latter include forest structure characteristics, vegetation types, slope, solar radiation, altitude, eastness, westness and variables extracted from the Aster imagery. Data exploration involved removing points locations falling outside the study area. The remaining samples were composed by 94 plots collected during Sep-Oct 2015 in Rwanda; these form a so called 'dataset 1' in this Thesis document. But also an additional dataset of 956 sample plots (collected during 2009-2010) was provided by Dr. Grueter Cyril in collaboration with the Dian Fossey Gorilla Fund International. These are referred to as 'dataset 2' in this Thesis. Both datasets contain the gorilla food species locations and abundance (dataset 1) or biomass (dataset 2), altitude, stem densities and corresponding vegetation types. The dataset 1 contains the tree canopy cover, dbh and tree height. In contrast, the dataset 2 contained the gorilla food species biomass obtained by harvesting, sun drying, weighing plants leaves & stems, followed by carrying out dimension analysis and solving phytometric regression equations (Grueter et al., 2013). On the other hand, the GIS and RS derived variables together with gorilla food species occurrence (derived from abundance or biomass) data were used for the Species Distribution Modelling (Figure 7).

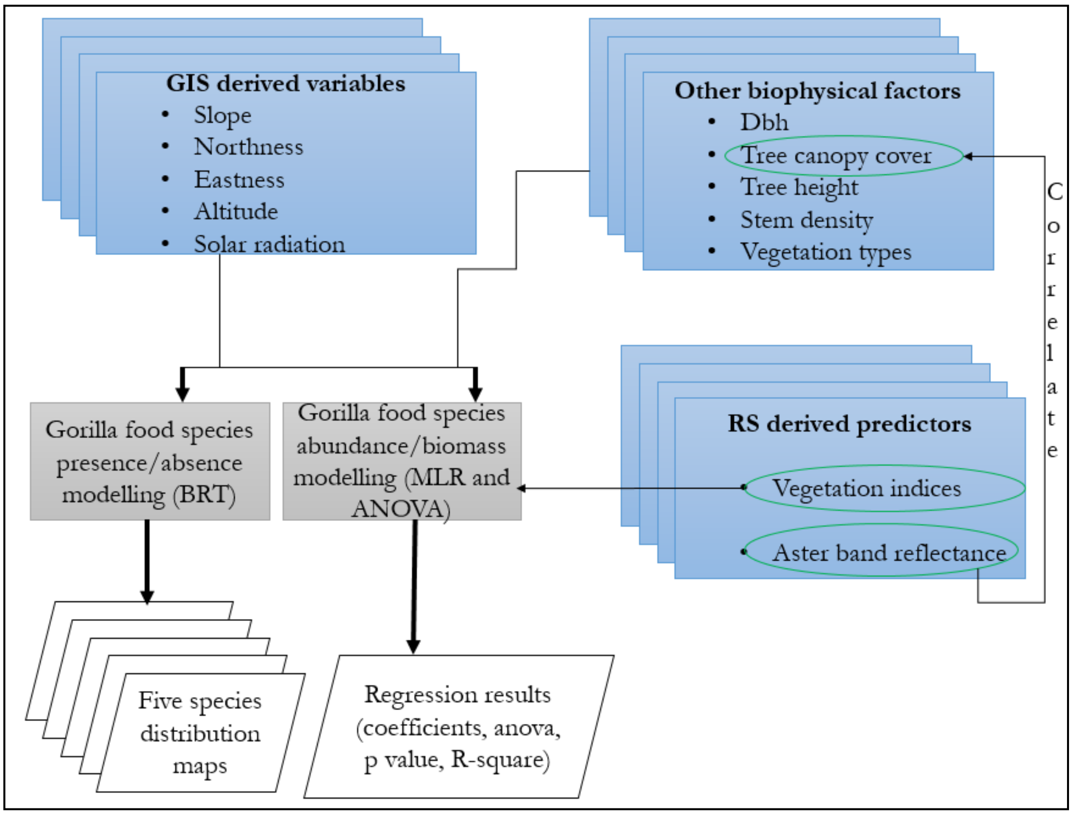


Figure 7: Summary of the methodological steps

Different software were used to achieve the objectives of this research. They are presented in the Table 6 below:

Table 6: Computer programs used in the study

Software	Use
• ERDAS IMAGINE 2015 • ArcGIS 10.3.1 • ENVI classic 5.3	Image processing and mapping
• Microsoft Excel 2013 • ArcGIS 10.3.1	Data preparation
• R software 3.1.2 • IBM SPSS statistics 23	Descriptive statistics Regression analysis
• Microsoft PowerPoint 2013 • ClickCharts Diagram & Flowchart software	Drawing flowcharts and diagrams
• Microsoft Word 2013	Thesis writing
• R software 3.1.2	Species distribution modelling (Boosted Regression Tree)

3.2. Study area: Virunga massif

3.2.1. Location

The Virunga massif lies between 1°20'0" to 1°40'0" latitude south and 29°20'0" to 29°40'0" longitude east (Figure 8) and covers approximately an area of 454 square kilometres. Home for mountain gorillas, the Virunga Massif consists of three parks shared among the countries namely Rwanda, Uganda and Democratic Republic of Congo (Plumptre et al., 2007). These parks are namely Parc National des Volcans (158.9 km^2) located in North-Western Rwanda, Parc National des Virunga (257 km^2) in Democratic Republic of Congo, and Mgahinga Gorilla National Park (39 km^2) in Uganda (REMA, 2011). From west to east, this area comprises six Volcanoes which are Mikeno (4437 m), Karisimbi (4507 m), Bisoke (3711 m), Sabyinyo (3634 m), Gahinga (3474 m) and Muhabura (4127 m). The DRC and Rwanda portions were created in 1925 as first African National Parks, while the Ugandan side park was established as a gorilla sanctuary in 1930 (Plumptre et al., 2007).

3.2.2. Biodiversity

Natural habitats as well as endemic fauna and flora makes the Virunga an area of continuous research and tourism attraction. Not only mountain gorillas, but also bird species, herbivores like buffaloes, bushbuck, black-fronted duiker and elephants inhabit the park. A total number of 480 mountain gorillas living in 36 groups with 14 solitary silverbacks has been recorded recently in the Virunga massif. Among these gorillas, 349 individuals found in 24 groups are habituated for research and tourism, while 101 individuals found in 12 groups are unhabituated (Gray et al., 2013). The findings from a recent survey on the biodiversity of the Virunga massif revealed 86 mammal species, among which 18 are endemic and 6 IUCN threatened species. 258 bird species among which 20 are endemic to the Albertine Rift and 4 are IUCN threatened species. 43 reptiles among which 7 are endemic to the Albertine Rift and none is categorized as IUCN threatened species. 47 amphibians among which 16 are endemic to the Albertine Rift and 9 IUCN threatened species. 878 plant species among which 124 are endemic to the Albertine Rift and 4 IUCN threatened species (Owiunji et al., 2005).

3.2.3. Soils

Virunga massif soils are generally fertile, but they vary from one park zone to another. Covered by a fine layer of humus and simple creeping flat roots on the steep slopes of the mountains, the soils in the Virunga are of volcanic origin; i.e. formed from volcanic ashes. They are in the category of Andosols and Andic soils with a black colour. In the wetlands, there is a phenomena of fine alluvial peat formation. Characterized by high moisture, rich in organic matter and high pH levels; the volcanic soils have a high permeability. The parent rock prevent water storage in the sub soil, which means that the region is not susceptible to soil erosion problems (Hitimana et al., 2006).

3.2.4. Human population

The Virunga massif is surrounded by mostly populated districts. For instance the Rwandan population has been increasing since 1978, and the current population density is estimated at 415 inhabitants per sq.km (MINECOFIN, 2012); particularly Musanze District adjacent to the Volcanoes National Park has a population density of 494 people per sq.km. The reason behind is that people have been attracted by volcanic fertile soils in the park surroundings, but also a cold climate which tolerates diseases like malaria carried by a mosquitoes preferring warm regions.

3.2.5. Climate

The annual rainfall in the Virunga region is approximately 2, 000 mm; with a distinct dry season starting from June to August, and heavy rainy season occurring in March throughout May. Intermediate seasons are the short rainy and dry season observed in the months of September-November and December-February respectively (Plumptre, 1991). Basically the rain falls in the area all the times of the year, but more heavy rain occurs from November to May (Fossey, 1974). In the Volcanoes National Park, the temperature drops with increased altitude whereas the wind speed increases in higher altitudes (Tuyisingize, 2010).

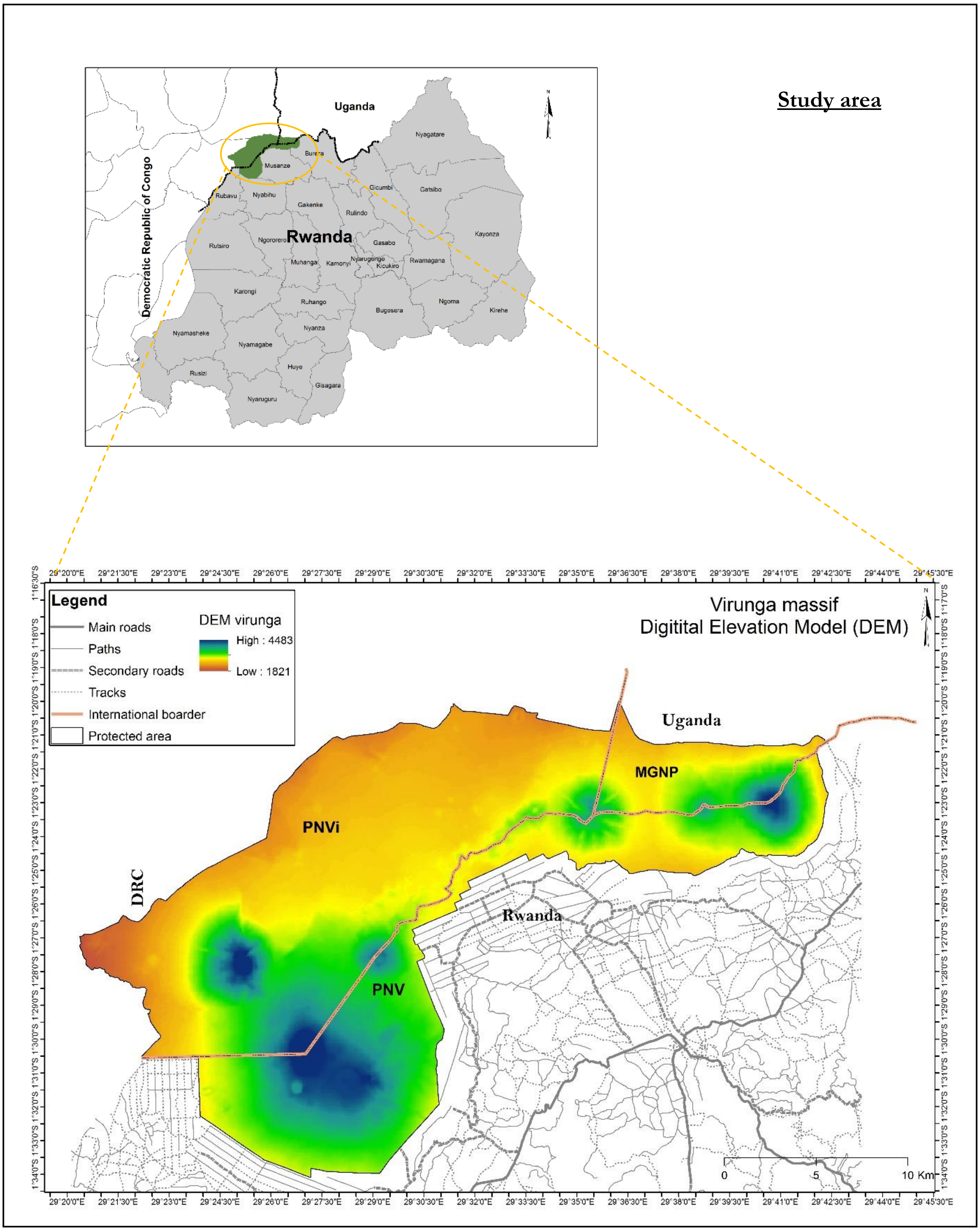


Figure 8: Location and elevation (in meters a.b.s.l) of the Virunga massif[a]

[a]MGNP: Mgahinga National Park; PNV: Parc National des Volcancs; PNVi: Parc National des Virunga; DRC: Democratic Republic of Congo.

3.3. Aster imagery pre-processing

The ITC Remote Sensing and GIS Lab ordered L1A Aster imageries (Table 7) and downloaded them from the Land Processed Distributed Active Archive Centre (http://LPDAAC.usgs.goc). Two scenes for each Aster band were mosaicked and band 4 to 9 were resampled to 15 m resolution. The layer stack was done in ERDAS IMAGINE 2015 followed by subset to get the imagery covering the study area.

Table 7: Aster imageries description

Sensor	Satellite	Date of acquisition	Spatial resolution	Swath width	Bands	Solar elevation angle	Image type
Aster	Terra	16/06/2006 21/02/2005	• 15 m for VNIR • 30 m for SWIR • 90 m for TIR	60 km	14 bands	56.733° 62.265°	AST14DMO

3.3.1. Converting DN values into TOA reflectance

The equations and other parameters necessary to convert the DN values into spectral radiance and then into TOA reflectance were available in the Aster user guide (Abrams et al., 2015; Ghulam, 2009; Gebreslasie et al., 2009):

DN values to spectral radiance:

$$L_{rad,j} = \left(DN_j - 1\right) \times UCC_j \qquad \text{Equation 1}$$

Where $L_{rad,j}$is Aster spectral radiance at the sensor's aperture measured in a wavelength j; j is the Aster band number; DN is the unitless DN values for an individual band j; UCC_j is the Unit Conversion Coefficient ($Wm^2sr^{-1}\mu m^{-1}$).

Spectral radiance to TOA reflectance:

$$\rho_{TOA,\lambda} = \frac{\pi * L_{rad,\lambda} * d^2}{E_{SUN,\lambda} * cos(\theta_s)} \qquad \text{Equation 2}$$

Where

ρ_{TOA}: Unitless planetary reflectance

L_{rad}: Spectral radiance at the sensor's aperture

d: Earth-sun distance in astronomical units

E_{SUN}: Mean solar exoatmospheric irradiances

λ: Wavelength, corresponds to the band number j

θ_s: Solar zenith angle (i.e. –solar elevation angle)

3.4. Fieldwork preparation

Before fieldwork, a 2008 aerial photograph of the Volcanoes National Park with 0.25 m resolution was provided by the Rwanda Natural Resources Authority (RNRA). It was prepared, saved as ECW format and uploaded into an IPAQ and other sub maps were printed to be used on field. The existing trails and the park boundary were provided by the Rwanda Development Board (RDB). Hence the Virunga vegetation map (Kayijamahe, 2008) was used to prepare a systematic sampling strategy (Figure 8) along trails but also covering all vegetation types.

3.5. Data collection

The field work was carried out in Rwanda during the period of September-October 2015. Following plot sizes used by Plumptre (1991) and Grueter et al. (2013) and taking into account that the study area is a dense forest; circular plots with 12.6 m of radius were used and the size of the entire plot was 500 m^2. The 30 meter measuring tape was used to know the boundary of the plot whereas the Clinometer was used to measure the slope and use a slope correction table (Appendix 1) to change the plot sizes where necessary. The forest characteristics variables (Table 8) were recorded on a data collection sheet (Appendix 2) as follow:

- 500 m^2 (entire plot): for trees (taller than 5 m); the tree dbh, height, canopy cover, dominant species and stem density.
- 5 m^2 subplot: for shrubs (height between 50 cm and 5 m); the names of the dominant species was written and the picture was taken for unknown species to be identified at the Karisoke herbarium.
- 1 m^2 subplot: for herbs (< 50 cm); the species abundance was recorded using the Braun-Blanquet approach. In the Braun-Blanquet method the plant cover is determined from estimates of vertical plant shoot-area projection as a percentage of quadrat area (Wikum & Shanholtzer, 1978). Scores representing the species abundance are then assigned. These are 0.5 (cover <1%), 1 (cover 1-5%), 2 (cover 6-25%), 3 (cover 26-50%), 4 (51-75%), 5 (76-100%) (Ellenberg & Mueller-Dombois, 1974).

The systematic sampling strategy was used where at each 300 m trail distance and 100 m buffer, a plot was taken alternatively at both sides of the trail (Figure 9). In total 94 plots were made in the east, centre and western part of the study area; in addition 956 sample plots (Figure 10) were provided by Dr. Grueter Cyril in collaboration with the Dian Fossey Gorilla Fund International (DFGFI). The number of samples taken in each vegetation type was quite different (Table 9).

Table 8: Forest variables recorded during fieldwork

Forest parameter	Measuring method/instrument	Units
• Location and altitude	GPS	meters
• Slope and aspect	Clinometer and Compass Suunto	percentage
• Tree canopy cover	Densiometer	percentage
• Dbh	Diameter tape	meters
• Plot size/radius	Measuring tape (30 m)	meters
• Tree height	Laser distance meter	meters
• Stem density	Counting the number of trees in the plot	integer
• Vegetation type	Based on literature, altitude and dominant species	text

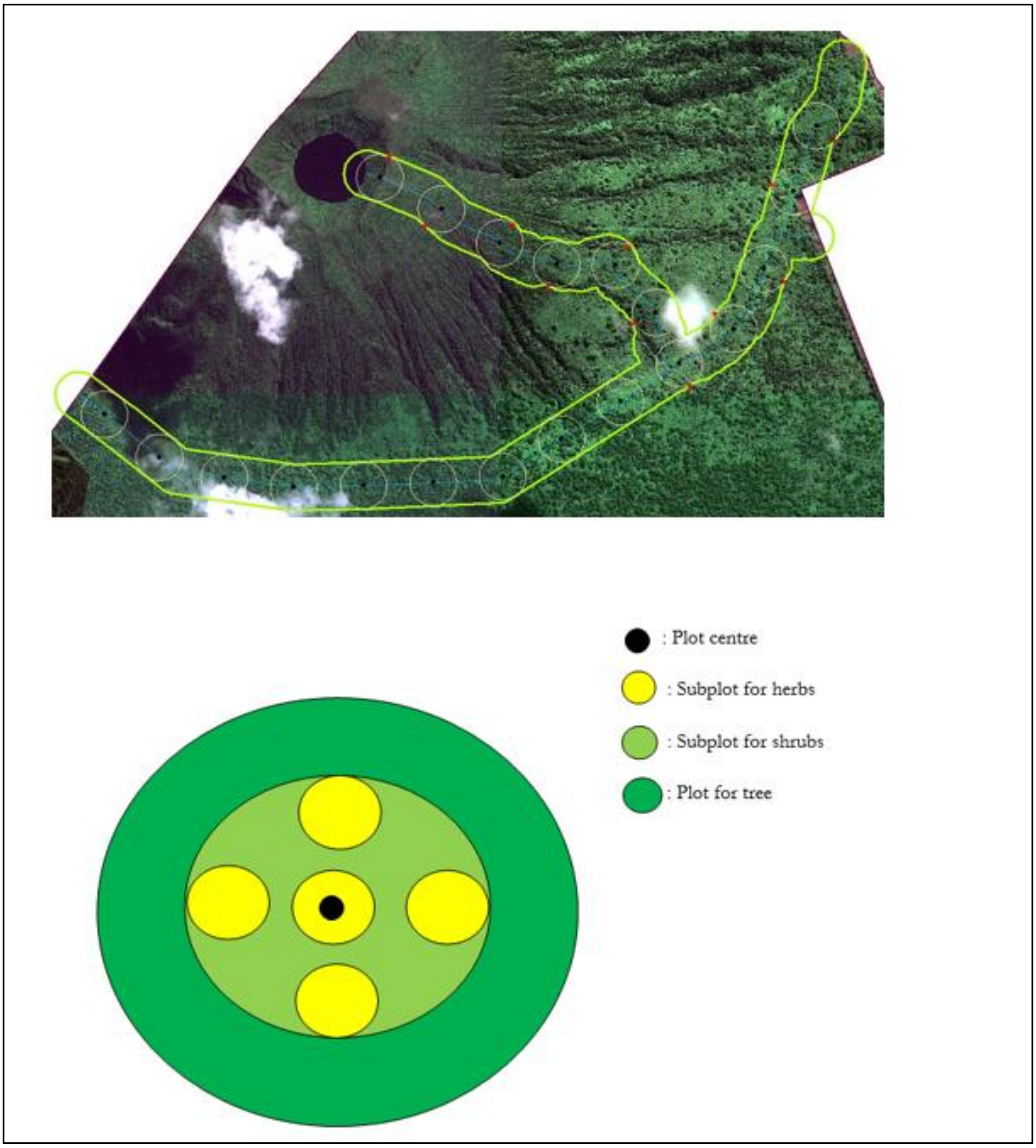


Figure 9: Systematic sampling (top) and data collection design in each plot (bottom)[a].
[a]The top aerial photo shows the 100 m buffer along trails (light green), point buffer around 300 m trail distance (beige) and sample locations (red). The trails are represented by blue lines and the dark dots are located at each 300 m of the trail.

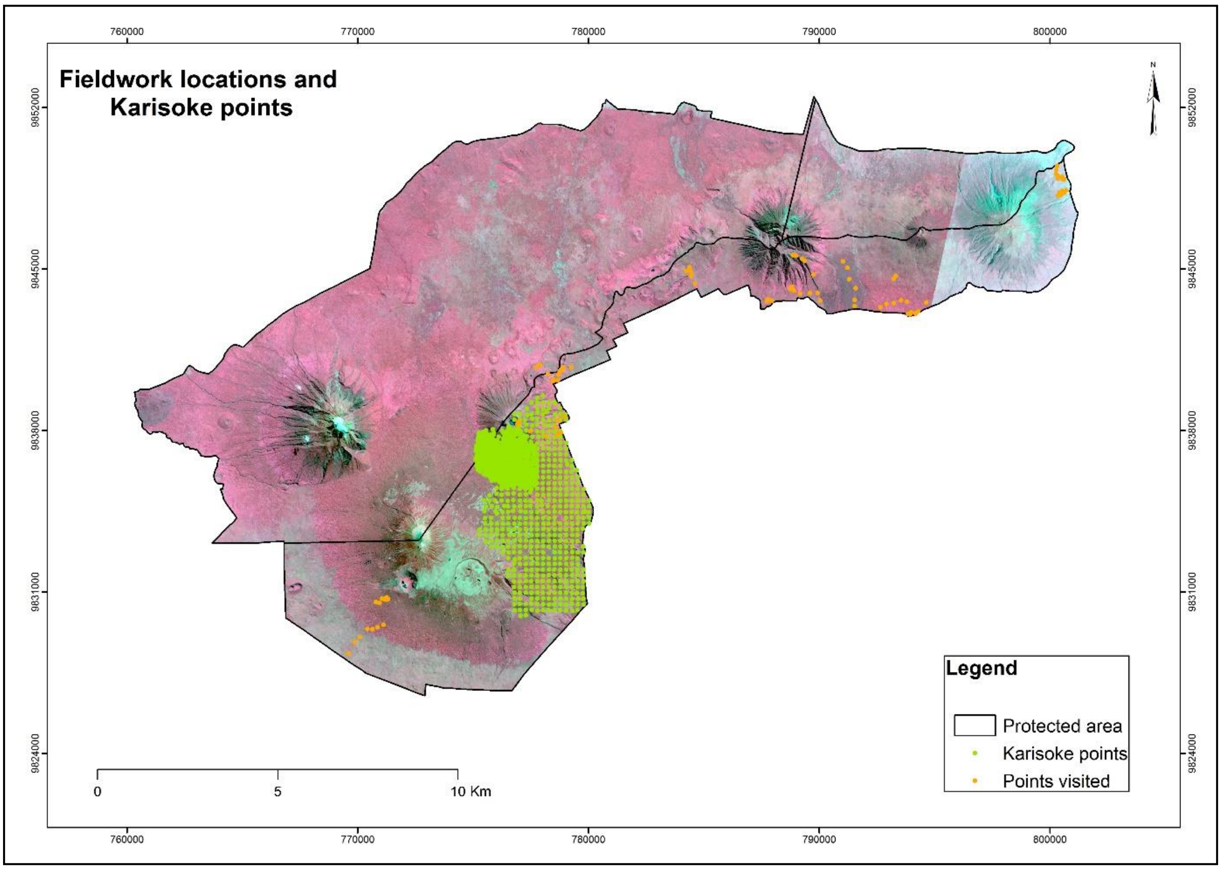


Figure 10: The 2006 Aster imagery covering the majority of the study area and the 2005 Aster on the upper right corner (Band combination 3, 2, 1 RGB)[a].
[a]From right to left the volcano names are: Muhabura, Gahinga, Sabyinyo, Bisoke, Karisimbi and Mikeno.

Table 9: Comparing the number of samples recorded in each vegetation type for the two datasets[a].

Vegetation types	**Number of samples**											
	Dataset 1						Dataset 2					
	tot	GAL	CRN	PLI	RUB	LAP	tot	GAL	CRN	PLI	RUB	LAP
Bamboo forest	14	1	3	0	0	1	33	6	0	0	5	12
Hagenia-Hyperic.	8	2	0	2	2	1	517	137	159	63	104	276
Herbaceous	13	3	1	6	1	2	58	27	16	14	7	26
Mixed forest	17	1	2	1	1	0	24	4	0	0	5	1
Neoboutonia	7	0	0	0	1	0	NA	NA	NA	NA	NA	NA
Brush ridge	6	3	2	0	6	0	164	61	45	19	77	53
Sub-alpine	4	0	0	0	1	0	80	27	7	0	63	0
Alpine	5	1	0	0	0	0	61	17	1	0	30	0
Meadow/savannah	11	5	1	0	3	0	20	4	0	0	2	0
Mimulopsis	9	2	0	0	1	0	NA	NA	NA	NA	NA	NA

[a]The tot column represents the overall number of samples taken in each vegetation type. The neighbour columns show how many samples (only presence) for individual gorilla food species were recorded in every single vegetation type. GAL: *Galium* spp., CRN: *Carduus nyassanus*, PLI: *Peucedanum linderi*, RUB: *Rubus* spp., LAP: *Laportea alatipes*. NA means that no samples were taken in that vegetation type.

3.6. Deriving vegetation types map from the Aster imagery

The supervised digital image classification using the Maximum Likelihood algorithm and accuracy assessment (Appendix 10) using ground truth data were performed in ERDAS IMAGINE 2015. The supervised classification is powerful because the operator defines the spectral characteristics of the classes by identifying sample areas (Richards & Jia, 2006). The vegetation types described in table 1 together with the Virunga massif vegetation map (Kayijamahe, 2008) were used to classify the study area into nine vegetation classes because the alpine and sub-alpine were combined into one class (Figure 11). The overall accuracy was 80.19 % with a kappa (K^) of 0.688 and 0.685 for the 2006 and 2005 imagery respectively.

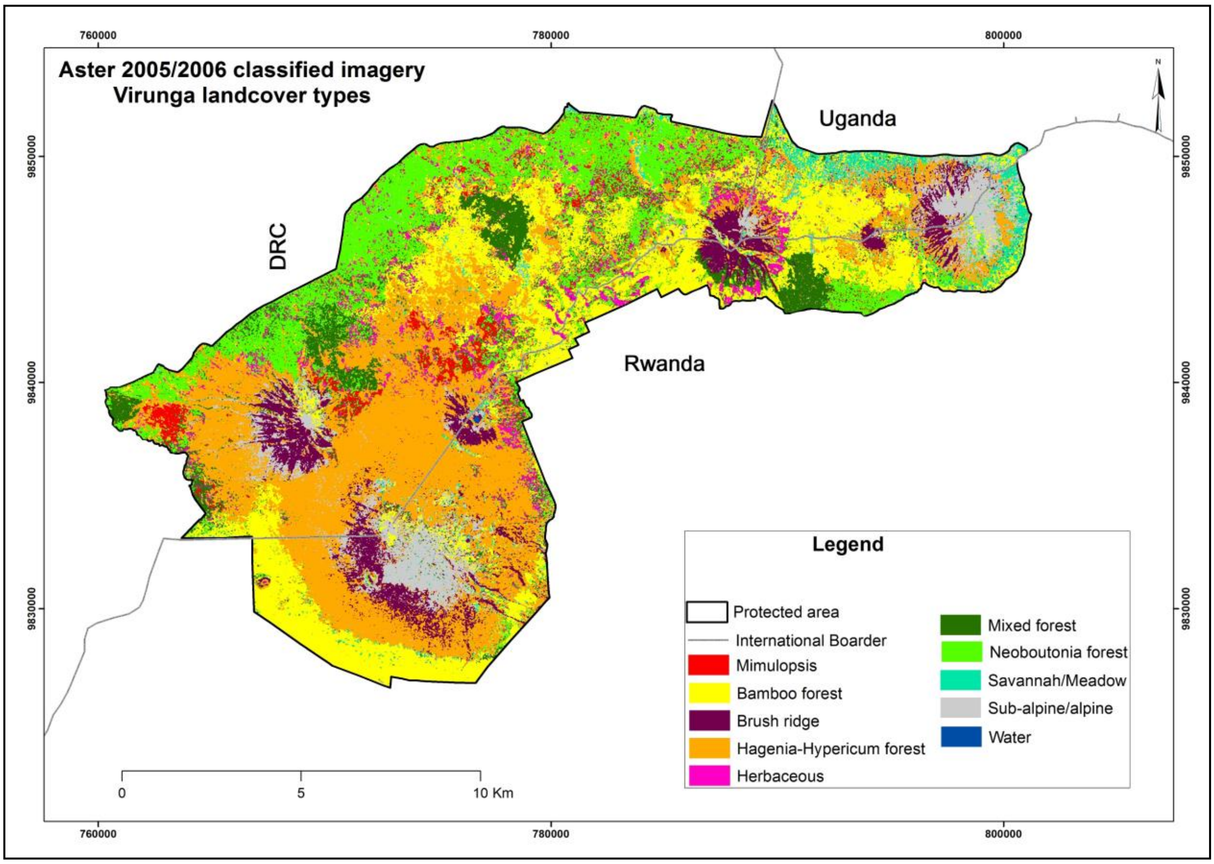


Figure 11: Vegetation types from the classified Aster 2006 and 2005 imageries

3.7. Data analysis

3.7.1. Significant difference in gorilla food abundance or biomass between vegetation types

The ANOVA test was used to determine if there is a significant difference in gorilla food species abundance or biomass between ten and eight vegetation types for dataset 1 and 2 respectively. The ANOVA test assumes that each of the population from which the sample comes is normally distributed and with the same variance. To check whether the samples come from a normally distributed population, both the Q-Q plots and Shapiro Wilk test were used whereas the Levene's test and looking at the standard deviations were used to check for variance homogeneity. However, the ANOVA test is robust against the normality assumption violation but when variances are equal (Field, 2009). The ANOVA result is accompanied by an F ratio-test degree of freedom, sum of squares and p value which are both indicators of the significant difference of means across more than two groups (Quinn & Keough, 2002). Therefore, because of uncertainty of knowing whether the population variances are equivalent; the Games-Howell post-hoc test was used for pairwise comparison when a significant difference was detected (Quinn & Keough, 2002). This analysis was not performed for *L. alatipes* for dataset 1 because of single observations (Table 9) made only in three vegetation types. The abundance greater than 40% was considered as higher while the biomass greater than 5 g/m^2 was considered as higher.

3.7.2. Multiple Linear Regression analysis

The explanatory variables were first tested for multicollinearity before considering them in a multiple linear regression model (MLR). On one hand, the response variable was the gorilla food species abundance or biomass while the explanatory variables included the forest characteristics together with topography and vegetation indices. On the other hand, the response variable was the forest canopy cover and the explanatory variable was the Aster band reflectance values. The multicollinearity exists when two or more predictors are strongly correlated ($r>0.5$); which results in providing redundant information and inflated standard errors of the estimates (Quinn & Keough, 2002). The Pearson correlation coefficient and the scatterplots of pairs of explanatory variables were used to detect multicollinearity. When it was detected, the Variance Inflation factor ($VIF=1/1-R^2$) was calculated which shows the strength of the relationship between each independent variable and all other covariates. Therefore the final variables to include in the regression model were selected based on VIF < 10 but also by taking into account the variable importance (Field, 2009 ; Quinn & Keough, 2002). Spatial autocorrelation occurs when random observations at pairs of locations separated by a certain distance (Legendre, 1993); which causes the inflated model accuracies (Veloz, 2009). The spatial autocorrelation was checked using the Moran'I (Fu et al., 2014), and calculating the average distance between point locations. The latter was 4.19 km which is bigger compared to the pixel resolution of the Aster imagery (15 m). Hence, none of the variables was removed because of spatial autocorrelation.

Fitted models

First, the forest characteristics parameters (Table 8) together with the topographic variables (Table 10) were used as indicators of mountain gorilla food species abundance or biomass. The aspect is a circular variable which was transformed into eastness and northness using respectively the sine and cosine functions to obtain a linear gradient (Piedallu & Gégout, 2008). For dataset 1, the gorilla food species collected using the Braun-Blanquet scores were converted into their corresponding percentage covers. For dataset 2 the forest characteristics parameters include the Hagenia and Hypericum tree densities.

Second, the field measured forest canopy cover was related to the Aster band reflectance values. This analysis was performed only on the dataset 1 because dataset 2 does not contain the forest canopy cover data. The major Aster imagery (2006) which was not covering the eastern part of the study area; was used for the extraction of the band reflectance values at each of the 80 locations (94 minus 14 falling in the eastern part).

The regression equation showed the relationship between the forest canopy cover and Aster band reflectance. The equation was then implemented in ERDAS IMAGINE 2015 using the model marker tool to create the forest canopy cover map of the Virunga massif.

Table 10: Topography related variables (see also Appendix 12)

Variable name	Units (range)	Source
DEM	Meters (30 m resolution)	Aster imagery
Slope	Degrees (0 to 90^o)	DEM
Solar radiation	Watts-hours/m^2	DEM
Aspect	Direction (west, east, south, north)	DEM
Eastness	West or east (-1 or +1)	Sine (aspect in absolute radians)
Northness	South or North (-1 or +1)	Cosine (aspect in absolute radians)

Lastly, using a linear regression model, the gorilla food species abundance or biomass was related to the vegetation indices. From the Aster band reflectance values, two vegetation indices (Table 11) values were calculated using the formulas below:

The Normalized Difference Vegetation Index:

$$NDVI = \frac{N - R}{N + R}$$

Equation 3

The two bands Enhanced Vegetation Index:

$$EVI2 = 2.5\frac{N - R}{N + 2.4R + 1}$$

Equation 4

where N and R in NDVI and EVI2 are the surface reflectance in the near-infrared and red Aster bands. The value 2.5 is the gain factor (G), 2.4 is the coefficient of the aerosol resistance term (C) while 1 is the soil adjustment factor (L).

The NDVI and EVI2 maps were created (Appendix 4) using the raster calculator tool in the ArcGIS 10.3 and then vegetation indices values for each of the locations in both dataset 1 and 2 were extracted and correlated to the gorilla food abundance or biomass.

Table 11: Selected vegetation indices

Vegetation index	Advantage	Reference
NDVI	• Robust and widely used • Its ratioing concept reduces illumination differences, cloud shadows, atmospheric attenuation and certain topographic variations	Tucker (1979)
EVI2	• Optimize vegetation signal • Improved sensitivity in high biomass regions • Performs well in heavy aerosols • Useful for sensors without a blue band	Jiang et al. (2008)

3.7.3. Mountain gorilla forage species presence modelling

The Boosted Regression Tree (BRT) model was selected for modelling the gorilla food species distribution because it performs better than traditional modelling techniques such as Generalized Linear Model (GLM) and Generalized Additive Models (GAM) (Albeare, 2009; Leathwick et al., 2006). Moreover, a slower learning rate (*lr*=0.001) was used in order to retain the stochastic component of model fitting, repeat 10 times the BRT model for each species and select the best model based on the performance accuracies (Elith et al., 2008). All model fitting, training, validation and plotting maps were done in R 3.1.2 software using the gbm package (R Development Core Team, 2014); but maps were saved and prepared in the ArcGIS 10.3.

a. Model inputs

First the gorilla food species abundance and biomass were converted into absence/presence data (i.e. where an abundance or biomass was recorded, there is obviously a presence). The total of 1036 observations (80 for dataset 1 and 956 for dataset 2) were used for each gorilla food species to extract the corresponding environmental conditions and then fit the BRT model (Figure 12). For each individual gorilla food species, the dataset was then split randomly by the R software 'runif ()' function; into 70% training samples and 30% validation samples (Liu et al., 2011. The response variable was binomial (family= 'bernoulli'; presence represented as '1' and absence as '0'). The same number of explanatory variables (Table 12) was used in order to allow comparison and variable reduction approach was not applied.

Table 12: Seven environmental variables selected for the BRT model input

Variable	Type	Name used in the model
Vegetation types	Categorical	vegty
Slope	Continuous	slope
Elevation	Continuous	dem
Solar radiation	Continuous	insolation
Canopy cover	Continuous	canopy
Eastness	Continuous	eastness
Northness	Continuous	northness

b. Model parameterization

Having the gbm package installed in the R software, the next step was to adjust the BRT model settings especially the tree complexity, the learning rate and the bag fraction. The rule of thumb described by Elith et al. (2008) suggests to first use the learning rate of 0.01, a tree complexity of 5 and a bag fraction of 0.5; but model settings should be applied based on the prevalence of the training data. A bag fraction of 0.5 (50%) controls the BRT model stochasticity and means that at each iteration, 50% of the data are drawn at random, without replacement, from the full training set. The bag fraction range of 0.5-0.75 has proven to give better model performance for presence-absence responses. Generally a smaller learning rate and larger number of trees is preferable although with smaller learning rates, the BRT model takes more time to fit. But they shrink the contribution of each tree more, and help the final model to reliably estimate the response (Elith et al., 2008). For the current study, a learning rate of 0.001, a tree complexity of 5 and a bag fraction of 0.5 were applied.

c. Model accuracy assessment

There is need to test for the predictions accuracy of presence-absence species distribution models. The main purpose consists of assessing the agreement between the presence-absence records and the predictions through examining the model discrimination capacity and reliability (Pearce & Ferrier, 2000). One widely used measure is the kappa statistics but which has been criticized to be dependent on threshold or prevalence (proportion of presence sample points in the whole sample). A new alternative measure of species distribution modelling (SDM) accuracy was proposed in ecology; that is the Area under the ROC Curve (AUC) which is independent of prevalence and the true skills statistic (TSS) (Allouche et al., 2006).

Like Kappa, TSS takes into account both omission and commission errors and success as a result of random guessing, ranging from -1 to +1, where +1 indicates perfect agreement and a value of -1 represents a perfect mismatch. The perfect agreement means that everything predicted as a presence is a presence and vice versa; while a perfect mismatch means that everything predicted as a presence is an absence and vice versa. TSS is not affected by either the prevalence or the size of the validation dataset (Allouche et al., 2006). To evaluate the presence-absence models, it is possible to compare the predictions with a validation dataset and then after construct a confusion matrix (Table 13). The latter depicts the sensitivity and specificity which are defined as the proportion of observed presence that are predicted as such and proportion of observed absences that are predicted as such respectively (Fielding & Bell, 1997).

Table 13: Presence-absence predictive accuracy error matrix*.

		Validation dataset	
		Presence	Absence
Model	Presence	*a*	*b*
	Absence	*c*	*d*

*a: number of cells for which presence was correctly predicted by the model (true positives); b: number of cells for which the species was not found but the model predicted presence (false positives); c: number of cells for which the species was found but the model predicted absence (false negatives); d: number of cells for which absence was correctly predicted by the model (true negatives) (Allouche, Tsoar, & Kadmon, 2006).

The sensitivity, specificity and TSS can be calculated as follow:

Sensitivity= $a/a+c$

Specificity=$d/b+d$

TSS=Sensitivity + specificity - 1

The ROC plot or AUC measures the ability of a model to discriminate between sites where a species is present versus those where a species is absent (Hanley & McNeil, 1982). A ROC plot is obtained by plotting the sensitivity (true positive rate) value on the *y* axis and their equivalent specificity (1-specificity or false positive rate) value on the *x* axis (Fielding & Bell, 1997). The AUC values range from 0 to 1, where a score of 1 indicates a perfect model discrimination capacity; a score of 0.5 implies predictive discrimination that is no better than a random guess and values less than 0.5 indicate a model performance worse than random (Elith et al., 2006). Pinkerton et al., (2010) described an AUC > 0.7 as a 'better-than-useful' degree of discrimination between presence and absence. For this study, the number of trees, both the TSS and AUC together with their standard deviations were used to evaluate the BRT model performance. The number of presence and absence in the validation dataset was different from one species to another (Appendix 14).
The percentage deviance explained by model was also suggested as an accuracy measure for the BRT model (De'ath, 2007). The MaxKappa was used to select the threshold (Liu et al., 2013) and go ahead with plotting the resulting maps. The percentage explained model deviance (pseudo R-square) is calculated as follow:

$$D^2 = 1 - (mean\ residual\ deviance)/(mean\ total\ deviance) \quad \text{Equation 5}$$

where D^2 is the percentage explained deviance

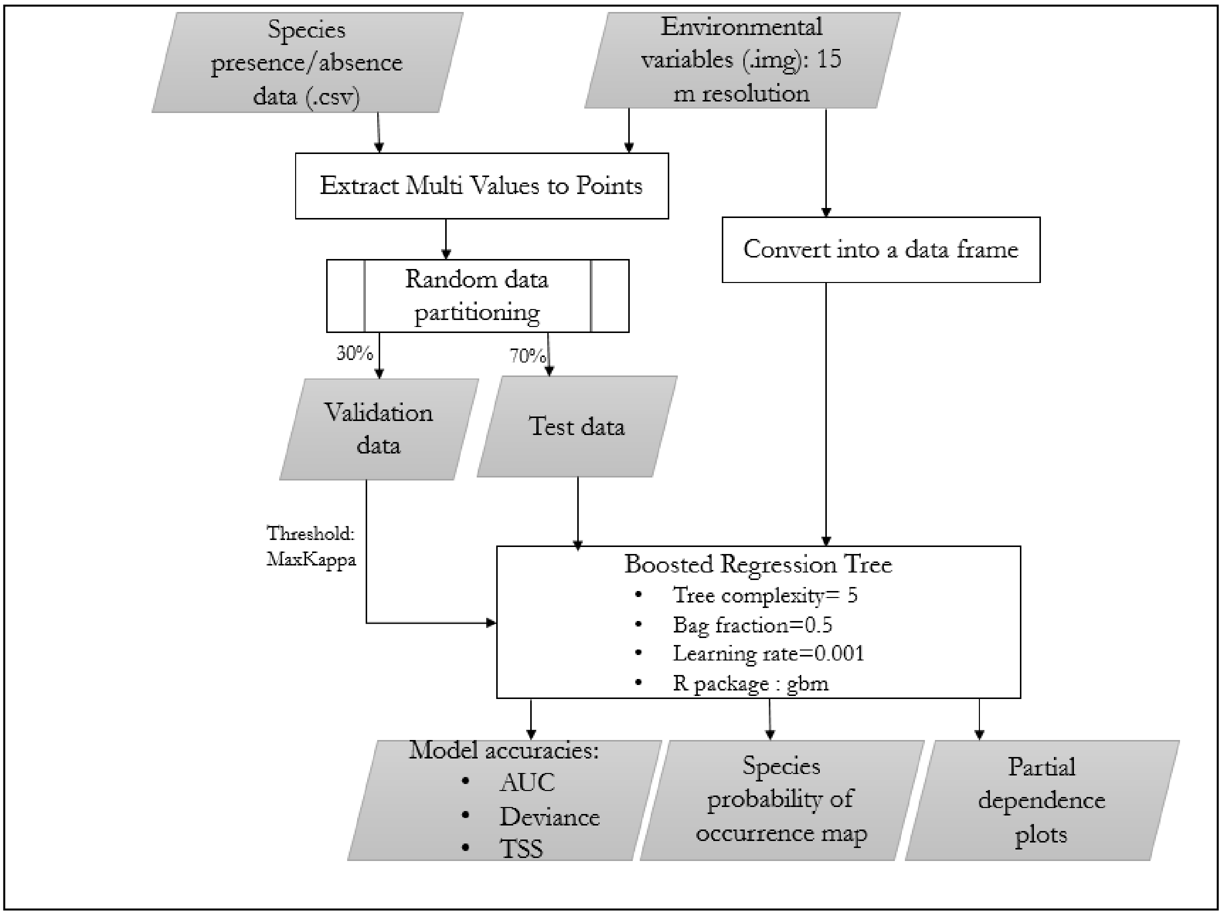


Figure 12: From BRT model fitting and validation to species distribution maps

d. Model output

The model training yields the number of trees and a graph, the mean total and residual deviance, the training AUC and cross-validation statistics. The model validation generates the relative variable importance, the partial dependence plots & fitted functions, the major interactions, the sensitivity, the specificity, the different thresholds, the AUC, the ROC plot and their respective standard deviation. The relative influence of predictor variables is measured based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split and averaged over all trees (Friedman, 2001). The partial dependence functions show the effect of a variable on the response after accounting for the average effects of all other variables in the model (Elith et al., 2008). Finally from the fitted BRT model, potential habitat distribution maps are plotted (Guisan & Zimmermann, 2000); those maps display the probabilities of species occurrence in a given study area.

4. RESULTS

In this chapter, the research findings based on objectives and research questions are presented. First, the significant difference in gorilla food species abundance or biomass between vegetation types. Second, the relationship between the gorilla food species abundance or biomass and forest structure together with topography variables as well as vegetation indices. Except for the species distribution modelling, the other results are presented per dataset and comparisons are made. The dataset 1 was collected during October 2015 fieldwork whereas dataset 2 was provided by Dr. Grueter Cyril in collaboration with the Dian Fossey Gorilla Fund International.

4.1. Significant difference in the abundance and biomass of gorilla food species between vegetation types

For dataset 1, none of the gorilla food species showed a statistically significant difference between vegetation types (p>0.05). Both the *Galium* spp. and *Rubus* spp. abundances were high in the Brush ridge and the two species could be found in the Mimulopsis and savannah/meadow vegetation types (Figure 13)

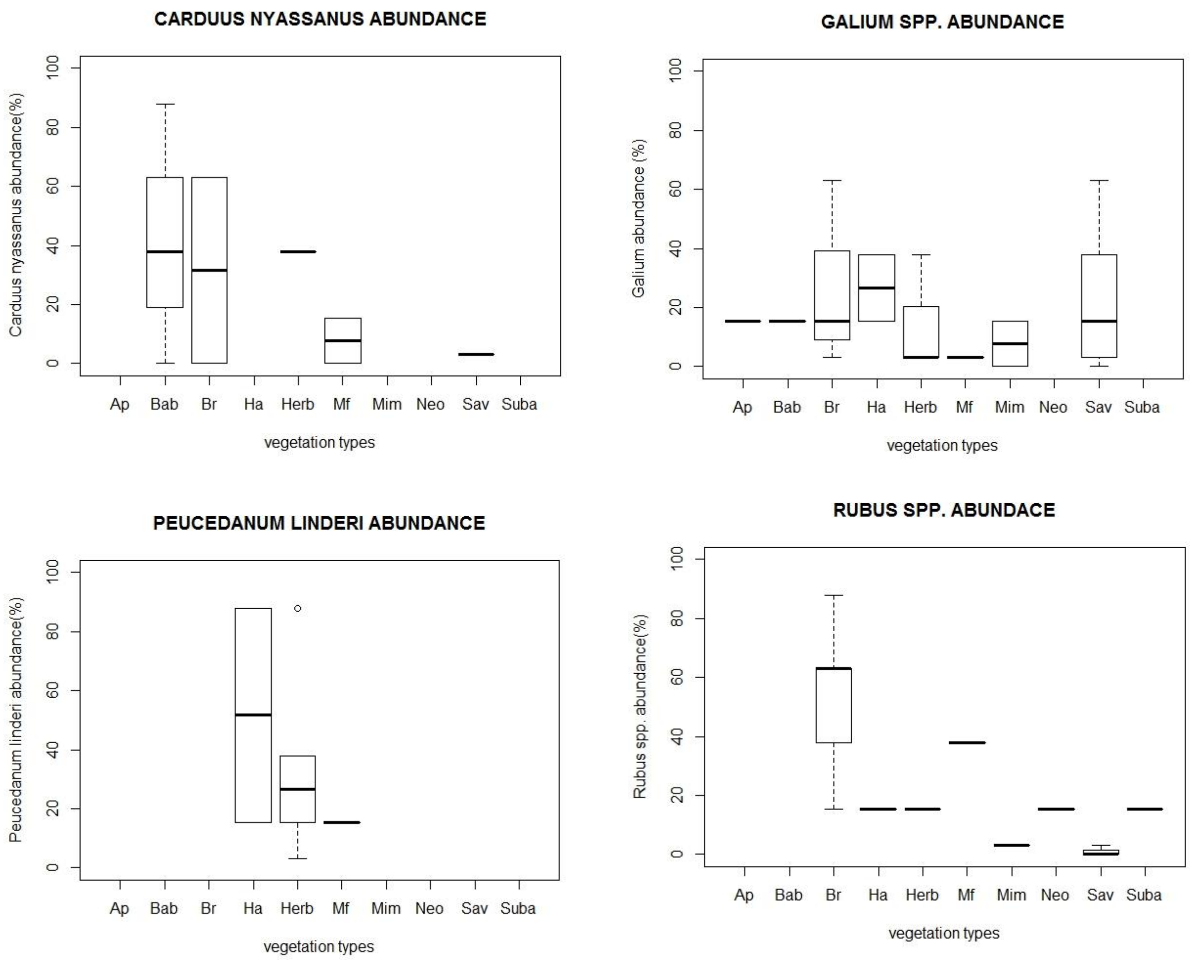


Figure 13: The abundance (%) in mountain gorilla food species between vegetation types[a] (dataset 1).

[a]Ap: Alpine; Bab: Bamboo; Br: Brush ridge; Ha: Hagenia-Hypericum; Herb: Herbaceous; Mf: Mixed forest; Mim: Mimulopsis; Neo: Neoboutonia; Sav: Savannah/Meadow; Suba: Sub-alpine vegetation types.

For both datasets, the *Galium* spp. higher abundances (>40%) and biomass (>5 g/m^2) were found in the brush ridge and sub-alpine vegetation types respectively (Figure 13 & 14). For the dataset 2, *Galium* spp. was significant ($F_{2,275}=2.164$; $p=0.03$); and the Games-Howell pairwise comparison showed a significant difference in *Galium* spp. biomass between the brush ridge and meadow and the sub-alpine and meadow.

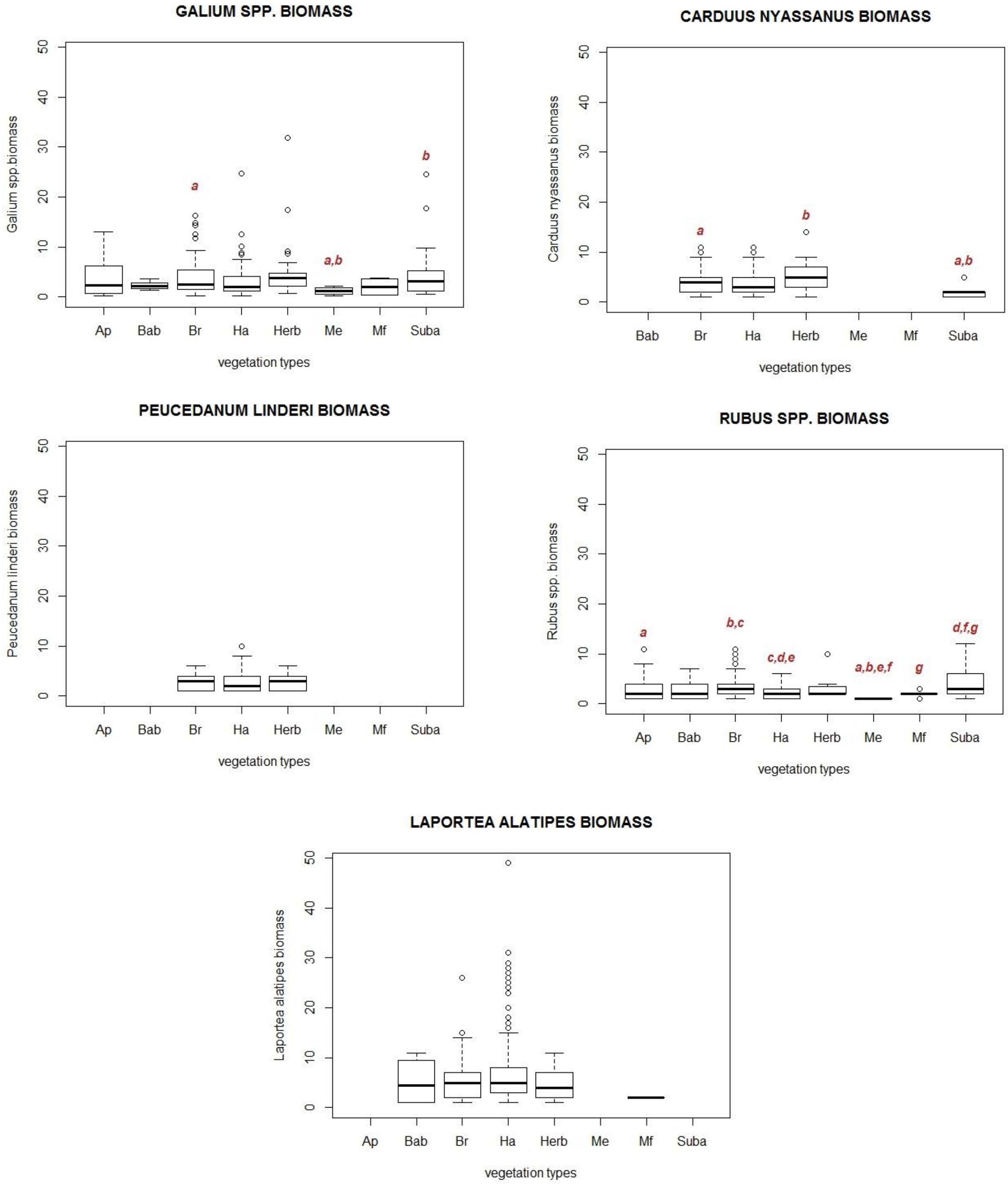


Figure 14: The mountain gorilla food species biomass (g/m^2) between vegetation types[a] (dataset 2).

[a]The boxplots without a label means there is no significant difference between pairs of vegetation types; while the labels 'a, b, c, d, e, f, g 'appear to mean a significant difference between groups of vegetation types. Ap: Alpine; Bab: Bamboo; Br: Brush ridge; Ha: Hagenia-Hypericum; Herb: Herbaceous; Mf: Mixed forest; Mim: Mimulopsis; Neo: Neoboutonia; Sav: Savannah/Meadow; Suba: Sub-alpine vegetation types.

Because there was only one observation of *C. nyassanus* in the alpine; the ANOVA was done except the alpine vegetation. The *C. nyassanus* biomass was significant different between vegetation types ($F_{3,223}=3.50$; p=0.01). The pairwise Games-Howell test showed a significant difference in *C. nyassanus* biomass between the brush ridge and the sub-alpine as well as the herbaceous and the sub-alpine vegetation types. The *Rubus* spp. biomass was significant ($F_{7,285}=6.91$; p<0.0001) and the Games-Howell pairwise comparison showed a significant difference between:

- The Hagenia-Hypericum and three vegetation types: brush ridge, sub-alpine, meadow.
- The mixed-forest and sub-alpine
- The brush ridge and meadow
- The sub-alpine and meadow
- The alpine and meadow

The *P. linderi* abundance and *L. alatipes* biomass were not statistically significant between vegetation types (Figure 14); but these two species were found in great quantity in the Hagenia-Hypericum and herbaceous vegetation zones (abundance > 40% and biomass > 5 g /m^2).

4.2. Relationship between the canopy cover and Aster imagery band reflectance values

The strong correlation between nine Aster band reflectance values was detected (r>0.5) from the scatterplots (Appendix 4) and calculated Pearson correlation coefficient (Appendix 9). The Variance Inflation Factor values showed a strong correlation between Aster band 4, band 5 with the rest of the bands (VIF>10). The two bands were not included in the stepwise regression.

Below is the final regression model after the stepwise procedure: the Aster bands 2, 3 and 7 were found to predict the forest canopy cover in the Virunga massif (Table 14). With 12.8% variation explained by the model, the canopy cover decreases with increased band 2 and 3 reflectance values, while the higher the band 7 reflectance value, the higher the canopy cover. In fact, the band 2 and 3 (Red and NIR) correspond to the high green vegetation cover. For the Virunga massif, the canopy cover was measured as zero in the herbaceous vegetation which showed higher reflectance value among other vegetation types. Therefore, the higher band 2 and 3 values, the lower the canopy cover relationship was found.

Table 14: Regression model results for the relationship between the forest canopy cover and Aster band reflectance values[a].

Coefficients	Estimate	Std. error	t value	Pr (>|t|)
Intercept	44.681	37.0924	1.205	0.2322
B2	-3.2042	1.2138	-2.64	**0.0101***
B3	-0.5035	0.2348	-2.145	**0.0353***
B7	4.8594	1.9577	2.482	**0.0153***

*: significance code: p<0.05
[a]($F_{3,74}=3.62$; $R^2=0.128$)

The canopy cover map of the Virunga shows low forest canopy on the side of Democratic Republic of Congo especially where the mimulopsis and secondary herbaceous formations are dominant. The higher canopy cover values were observed in the bamboo, the *Hagenia-Hypericum*, the *Neoboutonia*, the brush ridge and the mixed forest zones (Figure 15). There were differences in the measured canopy from the field and canopy cover predicted by the regression model (Figure 16).

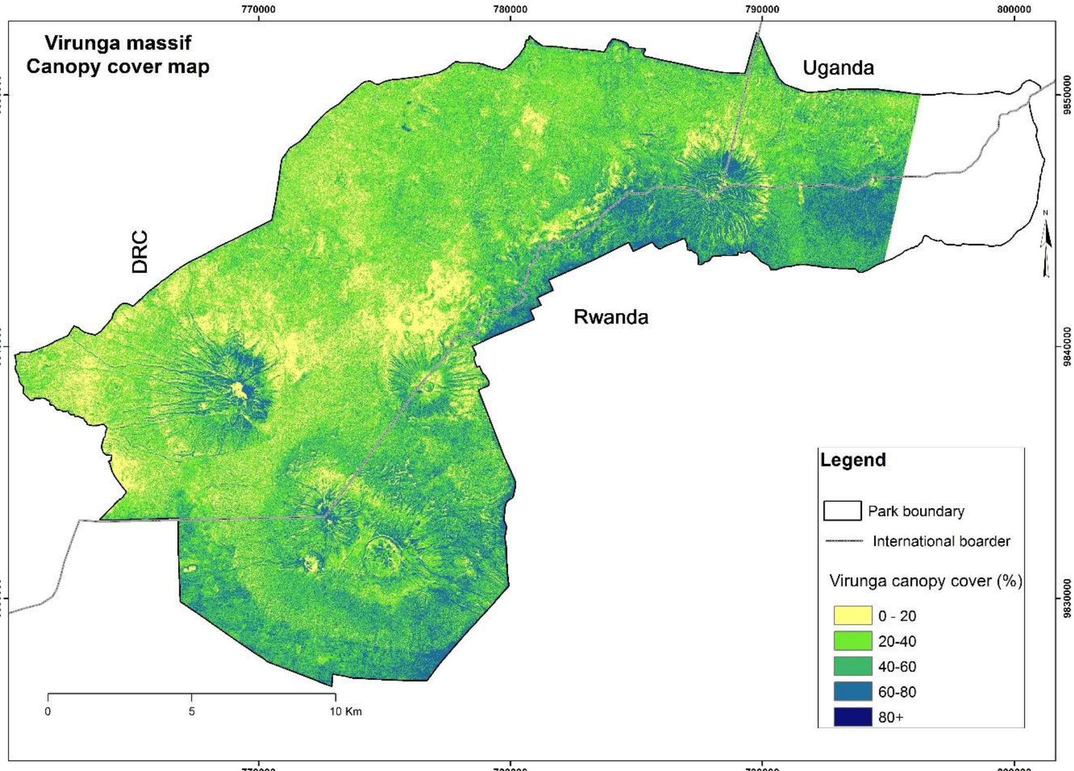


Figure 15: Forest canopy cover map of the Virunga protected area[a].

[a]The Eastern part of the park was not included in the canopy cover mapping because the 2006 Aster mainly used to extract the band reflectance does not cover that part.

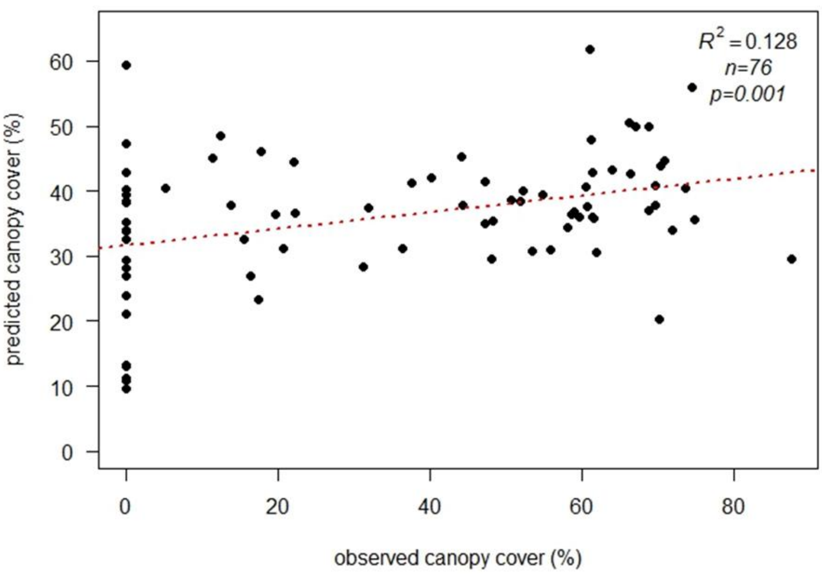


Figure 16: Relationship between the observed forest canopy cover and the predicted canopy cover at visited locations.

4.3. Relationship between forest structure characteristics, topography variables and gorilla food species abundance/biomass

In both datasets, there was a strong negative correlation between slope and solar radiation; in addition the height was positively correlated to both the tree canopy cover and diameter in dataset 1(Appendix 8). For dataset 1, the slope and solar radiation variables had a VIF value >10 and by calculating the VIF again without the solar radiation; the VIF values for the remaining variables was less than 10. Although the slope had a VIF >10 in the dataset 2, it was not eliminated so that the comparison between species and datasets can be made. Therefore, except the solar radiation, all other variables were used in the regression model for both datasets.

Table 15: Relationship between gorilla food species abundance/biomass (response variable) and forest structure characteristics together with topography variables (explanatory variables)[a].

	Dataset 1				Dataset 2			
Gorilla food species	Coefficients	Estimate	Std.error	t	Coefficients	Estimate	Std.error	t
GAL	intercept	**-20.98***	8.04	-2.60	intercept	-2.47^{ns}	1.44	-1.71
	elevation	**0.008***	0.00	3.04	elevation	**0.001***	0.00	2.18
	eastness	**-3.09***	1.52	-2.03	HAG	**-0.31***	0.16	-1.98
					slope	0.013^{ns}	0.007	1.75
CRN	intercept	**5.49***	2.00	2.73	intercept	**0.687***	0.145	4.71
	eastness	-3.32^{ns}	2.20	-1.51	eastness	**-0.27***	0.097	-2.80
	HT	-0.31^{ns}	0.21	-1.47	slope	**0.013***	0.005	2.30
					HAG	-0.17^{ns}	0.123	-1.43
PLI	intercept	**8.392***	2.77	3.021	intercept	**1.757***	0.550	3.19
	CC	-0.11^{ns}	0.06	-1.97	elevation	**-0.0004***	0.0001	-2.67
					northness	-0.083^{ns}	0.047	-1.74
RUB	intercept	**-59.00***	14.48	-4.07	intercept	**-8.891***	1.037	-8.57
	elevation	**0.021***	2.64	-1.59	elevation	**0.003***	0.0003	9.43
	HT	0.315^{ns}	2.51	1.39	HYR	**0.117***	0.043	2.68
	eastness	-3.68^{ns}	2.22	-1.46	eastness	0.164^{ns}	0.091	1.80
LAP	intercept	**0.497***	0.22	2.17	intercept	**19.276***	2.398	8.03
	northness	0.554^{ns}	0.31	1.76	elevation	**-0.005***	0.0007	-6.92
	eastness	-0.43^{ns}	0.30	-1.43	eastness	**-0.317***	0.211	-3.02
					HYR	**-0.317***	0.101	-3.13

[a]For both datasets, the significant coefficients ($p < 0.05$) retained after a stepwise regression model are presented with '*****' and ^{ns} for non-significant ones. CC: canopy cover; HT: tree height; HAG/HYR: Hagenia/Hypericum tree densities (number of stems/m^2). GAL, CRN, PLI, RUB, LAP stand for *Galium* spp., *Carduus nyassanus*, *Peucedanum linderi*, *Rubus* spp. and *Laportea alatipes* biomass or abundance.

For the dataset 1, only the *Galium* spp. and *Rubus* spp. regression equations were significant ($F_{1,67}$=13.65; R^2=0.19) and ($F_{4,73}$=8.059; R^2=0.30) with p< 0.05 respectively. In contrast to dataset 2, all regression models were significant (p< 0.05) but with a low variance explained by the model (R^2 < 0.12). The F-statistics were ($F_{4,951}$=3.99), ($F_{3,952}$=4.952), ($F_{3,952}$=4.007), ($F_{5,950}$=21.58), ($F_{3,952}$=22.36) for *Galium* spp., *C. nyassanus*, *P. linderi*, *Rubus* spp. and *L. alatipes* respectively.

In either datasets, the abundance or biomass of *Galium* spp. and *Rubus* spp. were found to increase in higher elevations while the *P. linderi* and *L. alatipes* biomass decrease with increased elevations. The reverse type of relationship exists for the *L. alatipes* and *P. linderi* where their biomass becomes low in higher elevations. Additionally, the *Rubus* spp. biomass rises in higher Hypericum tree densities while the *L. alatipes* biomass reduces with increased Hypericum tree densities (Table 15).

4.4. Relationship between mountain gorilla food species and three vegetation indices

For both datasets, there was a significant increase in *P. linderi* abundance or biomass in higher Normalized Difference Vegetation Index and two bands Vegetation Index values; while the *Rubus* spp. abundance or biomass showed a decrease in higher values of these two vegetation indices (Table 16). None of the vegetation indices predicted the abundance or biomass of *C. nyassanus* and *Galium* spp.

While in the dataset 2, the variance explained by the model was very low (<10%); 13% of the *Rubus* spp. abundance decrease in higher vegetation indices values was confirmed by the regression model in dataset 1. According to the two indices maps (Appendix 4), the higher vegetation indices values are observed in the Hagenia-Hypericum and herbaceous whereas the lower vegetation indices values occur in the brush ridge, sub-alpine and alpine vegetation zones. Thus the *Rubus* spp. abundance or biomass becomes high in the brush ridge and sub-alpine; while the *P. linderi* abundance or biomass was predicted high in the herbaceous and Hagenia-Hypericum zones of the Virunga massif.

Table 16: Relationship between mountain gorilla food species abundance or biomass with two vegetation indices[a].

	Dataset 1			**Dataset 2**		
Gorilla food species	Predictors	R^2	Equation	Predictors	R^2	Equation
GAL	NDVI	0.017^{ns}	y=1.842-2.143x	NDVI	0.0008^{ns}	y=-0.332+1.911x
	EVI2	0.024^{ns}	y=0.966-1.787x	EVI2	0.0000^{ns}	y=1.149-0.229x
CRN	NDVI	0.0004^{ns}	y=-0.057+0.363x	NDVI	0.0004^{ns}	y=0.124+1.078x
	EVI2	0.001^{ns}	y=0.386-0.457x	EVI2	0.002^{ns}	y=1.459-1.593x
PLI	NDVI	0.092*	y=-4.468+6.535x	NDVI	0.015*	y=-1.956+3.100x
	EVI2	0.13*	y=-1.834+5.548x	EVI2	0.013*	y=-0.428+2.012x
RUB	NDVI	0.042*	y=3.950-4.836x	NDVI	0.010*	y=4.545-4.988x
	EVI2	0.13*	y=2.721-5.965x	EVI2	0.017*	y=2.729-4.897x
LAP	NDVI	0.003^{ns}	y=-0.203+0.364x	NDVI	0.044*	y=-14.692+23.76x
	EVI2	0.001^{ns}	y=-0.012+0.194x	EVI2	0.017*	y=-1.164+10.23x

[a]The significant coefficients (p<0.05) are marked with '*' and '^{ns}' for the non-significant. GAL, CRN, PLI, RUB and LAP stand for *Galium* spp., *Carduus nyassanus*, *Peucedanum linderi*, *Rubus* spp. and *Laportea alatipes* abundance or biomass. NDVI: Normalized Difference Vegetation Index; EVI2: two bands Vegetation Index.

4.5. Spatial distribution of mountain gorilla food species in the Virunga massif

The Boosted Regression Tree (BRT) models for four species showed reasonable predictions (AUC > 0.70) while one species had an accuracy hardly better than random (AUC > 0.55) (Appendix 13). The BRT model for *Laportea alatipes* had the best discrimination ability as well as the higher deviance explained by the model; while the *Galium* spp. model had the lowest discrimination ability (Table 17).

Table 17: BRT model performance for each of the five mountain gorilla food species[a].

Gorilla food species	Nt	MaxKappa	Sens.	Sens. sd.	Spec.	Spec. sd.	TSS	AUC	AUC sd.	D^2
RUB	3250	0.37	0.61	0.047	0.84	0.024	0.45	0.78	0.028	0.31
GAL	2150	0.34	0.45	0.051	0.79	0.026	0.24	0.65	0.034	0.17
CRN	2050	0.30	0.53	0.056	0.83	0.023	0.36	0.77	0.028	0.21
PLI	1300	0.13	0.53	0.090	0.81	0.023	0.34	0.72	0.051	0.18
LAP	2700	0.41	0.71	0.043	0.76	0.030	0.47	0.80	0.025	0.32

[a]Nt: number of trees, sd.: standard deviation, D^2: percentage explained deviance (pseudo-R-square). RUB: *Rubus* spp., GAL: *Galium* spp., CRN: *Carduus nyassanus*, PLI: *Peucedanum linderi*, LAP: *Laportea alatipes*. AUC: (validation) Area Under the Curve. Sens.: sensitivity; Spec.: specificity.

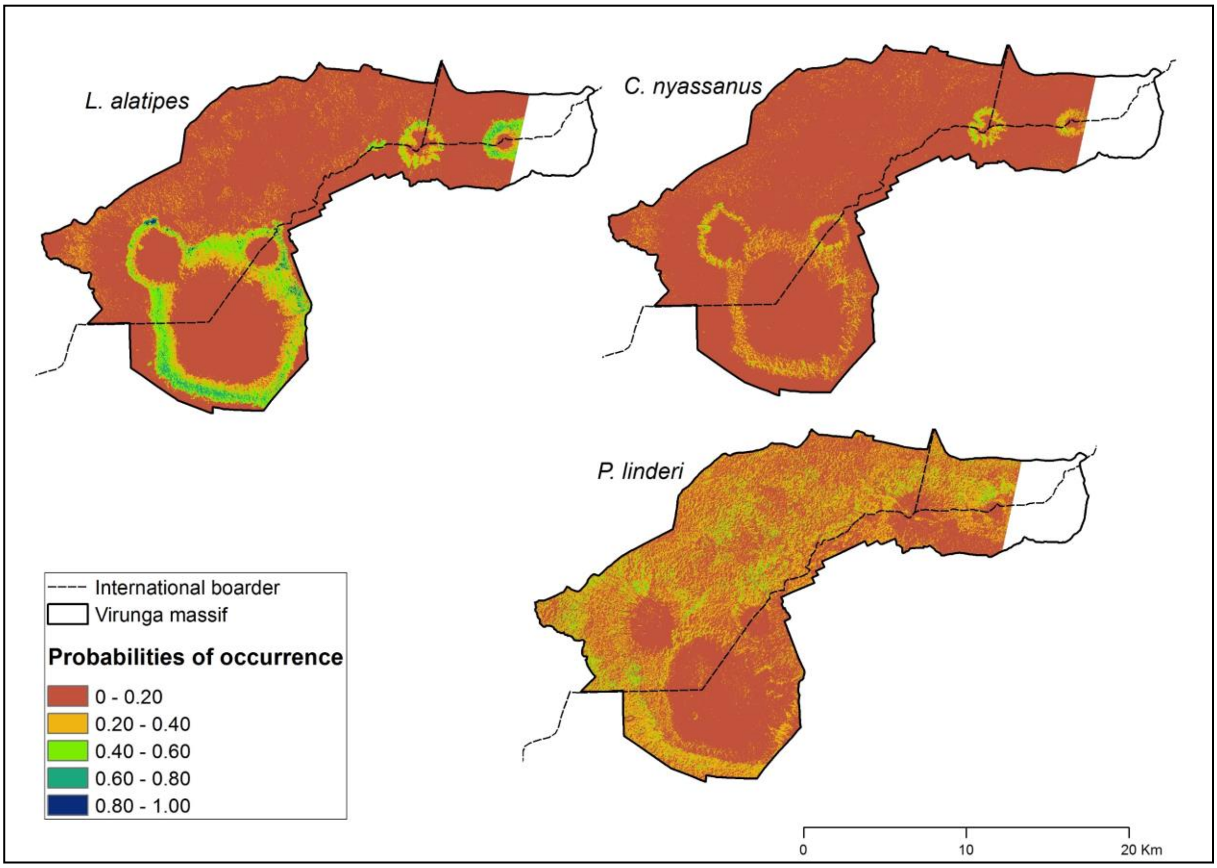


Figure 17: Spatial distribution of *Laportea alatipes*, *Carduus nyassanus* and *Peucedanum linderi* in the Virunga massif

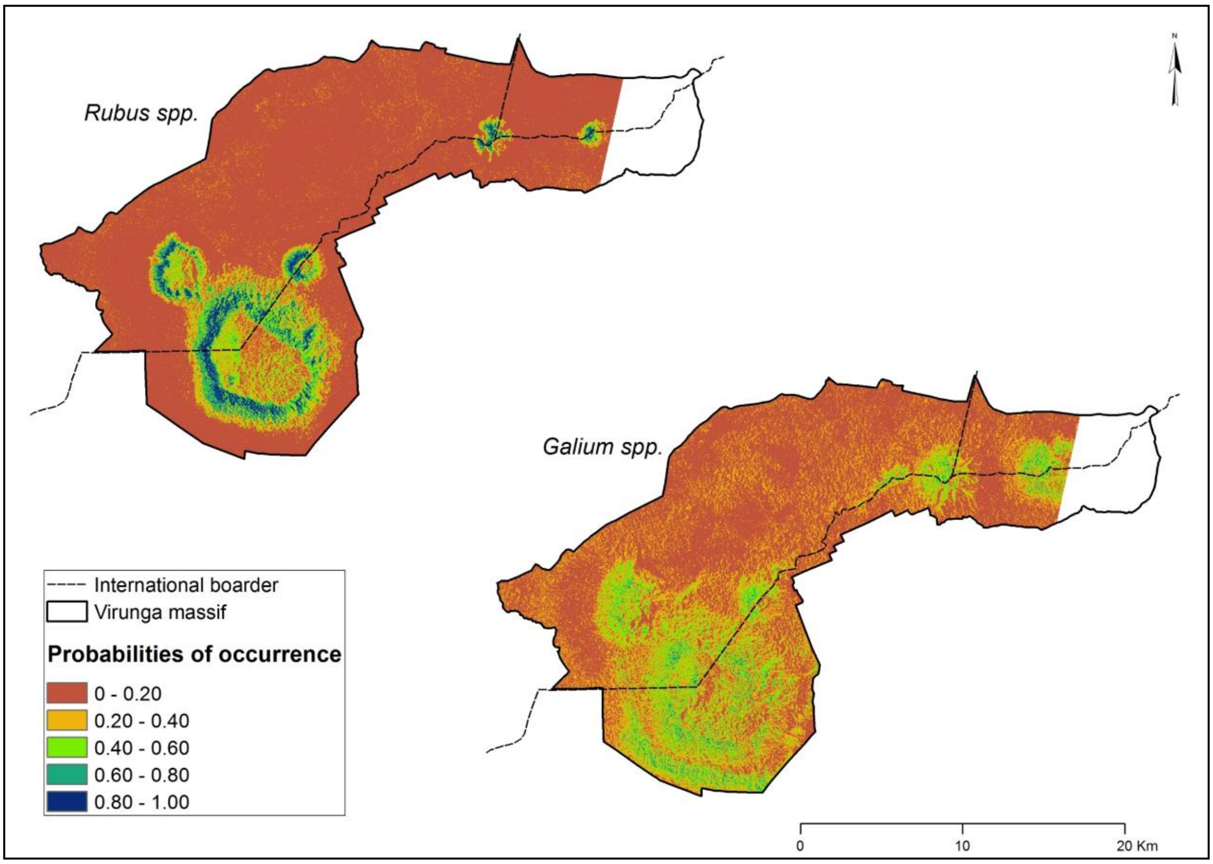


Figure 18: Spatial distribution of *Rubus* spp. and *Galium* spp. in the Virunga massif

The predictor relative importance for each species (Figure 19 & 20) showed that elevation and eastness were among the three most important variables for predicting the occurrence of each of the five gorilla food species (Table 18).

Table 18: The three most important predictors for the occurrence of the gorilla food species

Gorilla food species	Three most important predictor variables
Rubus spp.	Elevation Eastness Slope
Galium spp.	Eastness Elevation Solar radiation
Carduus nyassanus	Elevation Eastness Northness
Peucedanum linderi	Eastness Elevation Forest canopy cover
Laportea alatipes	Elevation Vegetation types Eastness

Both *L. alatipes* and *P. linderi* have very low probabilities of occurrence in higher elevations especially on volcanoes summits (3600 m - 4500 m). The BRT model predicted higher occurrences of *Galium* spp. and *Rubus* spp. in higher altitudes coupled with western facing slopes and their presences are observed even on the volcanoes peak (Figure 19 & 20). The optimal suitable elevation range for *Rubus* spp. was between 3200 m-3500 m with probabilities of occurrence in elevations less than 3000 m very close to zero.

The *C. nyassanus* was found to occur in the middle altitudes of the Virunga protected area. Its probability of occurrence starts increasing from the altitude of 2800 m and reaches the maximum at 3200 m where it then drops. The five species did not show any remarkable variation in their occurrences with respect to the northness; but they prefer western facing slopes. The western part of the Virunga massif (Karisimbi and Mikeno mountains) is the most suitable for the growth of *Rubus* spp., *Galium* spp. (Figure 18). Except the *P. linderi*, the four other species have very low to absent probabilities of occurrence in lower altitudes (< 2600 m) of the Parc National des Virunga in DRC (Figure 17 & 18).

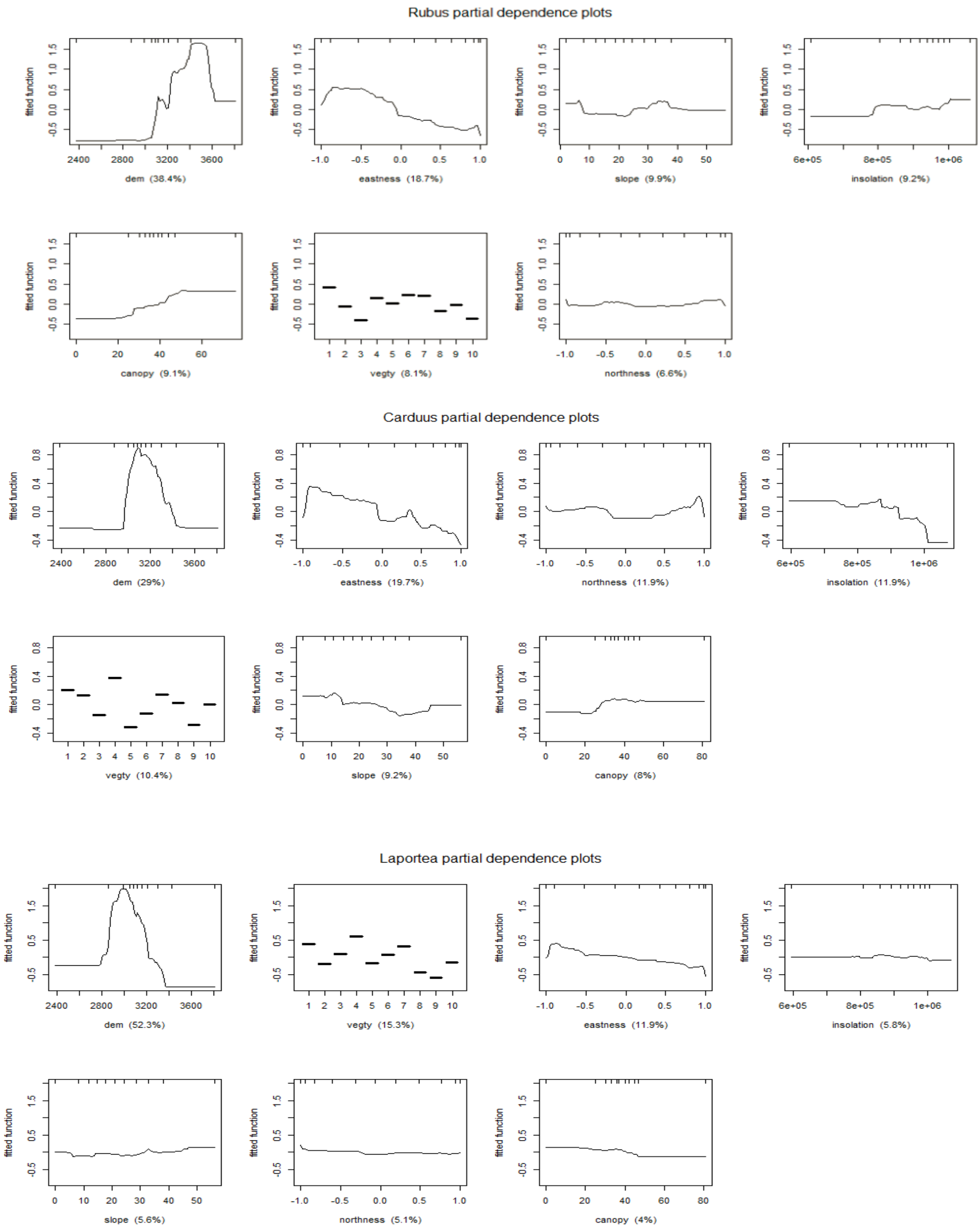


Figure 19: Partial dependence plots for three mountain gorilla food species[a].

[a]For each species, the variables are ordered by decreasing relative importance (percentage written between brackets).

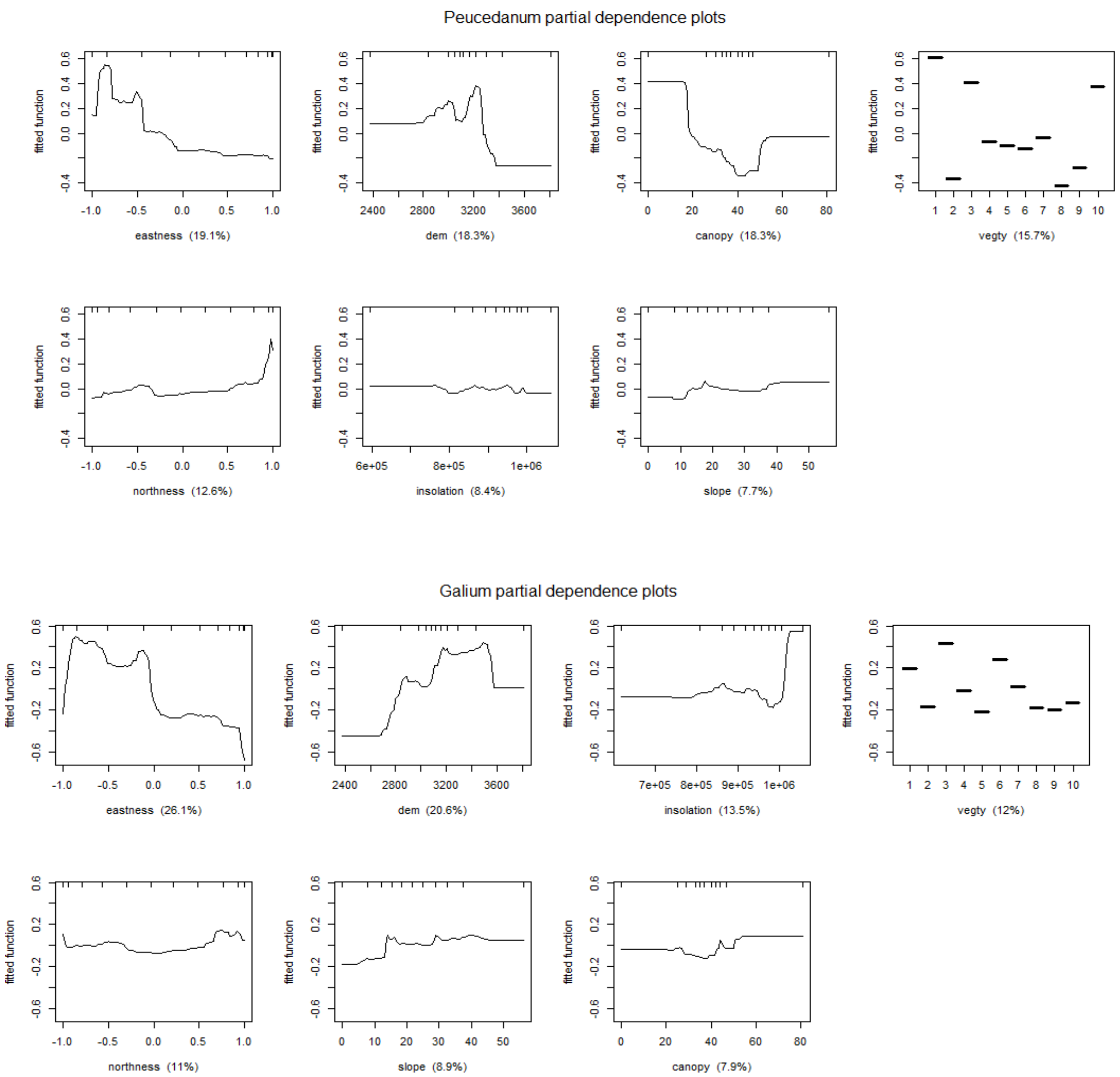


Figure 20: *Galium* spp. and *P. linderi* partial dependence plots with the eastness as the most important predictor

5. DISCUSSION

The aim of this study was to assess the usefulness of Remote Sensing (RS) approaches for mapping the spatial distribution of five most preferred mountain gorilla food species. Since these gorilla food species are mainly below-canopy plant species and cannot directly be detected by the Aster imagery; indirect mapping techniques have been applied. The relationship between each of the five gorilla food species abundance or biomass with environmental variables was determined. The latter included vegetation types, forest structure & topography variables and vegetation indices. Variables that could be mapped with RS were then used for the spatial distribution modelling of each of the gorilla food species. In this chapter, the focus is on the reflection on the choice of biophysical explanatory variables, next, the model accuracies and finally the individual species abundance and occurrence.

5.1. Biophysical variables selection

The gorilla food species are abundant and perennially available in the gorilla habitat (Watts, 1998a). The latter is a tropical rainforest where vegetation structure and composition vary with altitude (Owiunji et al., 2005). Vegetation types with open canopies allow the growth of gorilla food species while closed canopies vegetation types have poor understory (Weber & Vedder, 1983). The amount of light, the soil moisture or composition are direct factors determining the abundance of plant species in the Virunga massif (Plumptre, 1991). Dondeyne et al.(1993) found that typical volcanic soils (Andosols) can harbour higher abundances of the gorilla food species. Light intensity requires daily measurements either using the hemispherical canopy photography technique or photometers (Joshi et al., 2006). Other variables impacting on Virunga plant species could be the trampling effect by large herbivores; but Plumptre (1991) found that the plant damage by herbivores had little effect on the food-plants biomass. In addition, it has to be tested by a model if a certain variable really has an important contribution to predict species abundance and occurrence (Plumptre, 1996). However, the soil, light intensity and grazing pressure variables were lacking for the current study. From the Aster DEM, topography variables which have an influence to the light intensity, soil moisture, temperature, rainfall and wind can be extracted. Therefore, vegetation types and indices, forest structure together with topography variables have been used in this study to predict the abundance or biomass and spatial distribution of gorilla food species. Selecting the most important variable was done through calculating the Variance Inflation Factor, the stepwise regressions as well as the researcher's own judgement.

5.2. Model accuracies

One way Analysis of Variance (ANOVA) was used for determining the statistical significant difference in gorilla food species abundance or biomass between vegetation types. The one way ANOVA F-test assesses the overall fit of the model to the data and it appears to be robust to moderate violations of normality; but with equal sample sizes (Quinn & Keough, 2002). Hence, the Games-Howell test was used for pairwise comparison. For dataset 1, none of the gorilla food species showed a significant difference in abundance between vegetation types; while for dataset 2, there was a significant difference in three of the species biomass between vegetation types. One reason for non-statistically significant results can be attributed to the small sample size or few species observations in each of the vegetation types.

The stepwise multiple linear regression was used for predicting the Virunga forest canopy from the Aster imagery band reflectance values. The resulting regression model was significant ($p<0.05$) but with only 12.8% of forest canopy cover explained by Aster band reflectance values. In addition, there were differences in the predicted forest canopy cover and the observed canopy cover from field (Figure 15). These differences highlight the uncertainties and limitations associated with mapping the forest canopy cover using remote sensing. For instance using the Landsat 8 imagery, the multiple linear regression method for canopy density mapping was found to perform worse compared to the artificial neural network, the forest canopy density mapper or maximum likelihood classification (Joshi et al., 2006). Moreover, the Aster imagery used for this study was acquired June 2006 and data collection was carried out October 2015. Hence, both the 15 m Aster image resolution together with differences in field data and image acquisition date are possible reasons for a lower variance explained by the forest canopy cover model in the current study. Looking at the vegetation types map (figure 10), the closed canopies (>60%) were predicted in the bamboo, mixed forest, Hagenia-Hypericum forest, *Neoboutonia* forest and the brush ridge; while the open or zero canopies (< 20%) were found in the herbaceous, Mimulopsis and alpine vegetation zones (Figure 15). The high and low canopy covers in the similar vegetation types were also obtained through vegetation cover estimation in the Volcanoes National Park (Tuyisingize, 2010); one of the protected areas composing the Virunga massif. The zero forest canopy cover in secondary herbaceous formations (Appendix 11) can be indicators of forest disturbances (Dondeyne et al., 1993). The negative relationship between the forest canopy cover and the near-infrared resulted from the field zero canopy cover recorded in the herbaceous vegetation type. These consist of dense tall herbs with no tree cover; and they were found to have a high reflectance compared to the bamboo and Hagenia-Hypericum forests in the Virunga massif (Appendix 15).

The stepwise multiple linear regression was used to predict the gorilla food species abundance (dataset 1) or biomass (dataset 2) from both the forest characteristics parameters and the topography variables. For dataset 2, all five species regressions were statistically significant ($p<0.05$) while two species showed a significant regression for dataset 1. For both datasets, *Rubus* spp. and *Galium* spp. regression models showed 30% and 19% of these species variation in abundance with regard to forest structure and topography variables; the rest of the model variance explained is lower than 12%. In addition, for dataset 1, none of the field measured forest structure variables (tree diameter, forest canopy cover, tree height) were significant. The possible reason is the small observations made in forest areas for each of the five species (Table 9).

The linear regression model was used to test how vegetation indices can predict the abundance or biomass of gorilla food species. Non-significant coefficient were observed for three species in dataset 1 and two species for dataset 2 (Table 16); and with very low variance explained by the model. This result shows that both the NDVI and EVI2 have a poor prediction accuracy for the abundance or biomass for *L. alatipes*, *C. nyassanus* and *Galium* spp. For dataset 1, the EVI2 could explain 13% of the abundance of *Rubus* spp. and *P. linderi* than NDVI. Gu et al. (2013) confirms that NDVI is a poor predictor of forage abundance in higher biomass landscapes.

The Boosted Regression Tree was used for mapping the spatial distribution of each of the five gorilla food species. The Area Under the ROC Curve (AUC) as one the measure of model accuracy. Four of the species had an AUC > 0.70 which shows reasonable predictions while one species had an AUC=0.65 which shows predictions hardly greater than random (Duque-Lazo et al., 2016). While the stepwise variable reduction improves the species distribution model accuracy (Van Gils et al., 2014); this approach was not used in this study because the same number of variables was kept in order to allow comparison between species. Additionally, none of the variables in this study was removed because of the spatial autocorrelation.

5.3. Gorilla food species abundance or biomass and spatial distribution

The five gorilla food species behave differently towards environmental conditions. The paragraphs below describe the abundance and occurrence of each of these species in the Virunga massif. The 'abundance' term is used for dataset 1 while 'biomass' is used for dataset 2.

5.3.1. *Rubus* spp. (blackberry)

The *Rubus* spp. biomass was significant between seven vegetation groups. The higher biomass (> 5 g/ m^2) of this species was found in the brush ridge, sub-alpine and alpine vegetation zones. These results are in accordance with Plumptre (1991) who measured 0.13 g/m^2; 0.11g/m^2; and 0.34 g/m^2 biomass of *Rubus* spp. in the brush ridge, the sub-alpine and alpine vegetation zones respectively. The stepwise multiple linear regression showed that *Rubus* spp. abundance or biomass significantly increased in higher altitude. The gorilla food quantity determines the area occupation by mountain gorillas (Watts, 1985); therefore this results partly explains why lower altitudes (< 2600 m) of the Virunga side in Democratic Republic of Congo are rarely used by mountain gorillas. The *Rubus* spp. BRT model (AUC=0.78) predicted higher probabilities of occurrence (> 0.60) of *Rubus* spp. in the elevation ranges of 3200 m- 3500 m and western-facing slopes. Sternberg & Shoshany (2001) highlight that higher resources availability especially through soil moisture and competition for light explain the abundance of vegetation on humid slopes. Therefore, for the case of the Virunga massif, with western-facing slopes more humid than eastern-facing slopes; the vegetation biomass is higher on the western- facing slopes compared to eastern-facing slopes. The *Rubus* spp. abundance or biomass was found to decline with higher NDVI and EVI2 values.

5.3.2. *Galium* spp.

The *Galium* spp. biomass was significantly different between two vegetation groups. The *Galium* spp. is available in all vegetation types expect the Neoboutonia (figure 13 & 14). The higher biomass of *Galium* spp. was (>5 g/m^2) was found in the herbaceous, Hagenia-Hypericum, brush-ridge, sub-alpine and alpine vegetation zones. Plumptre (1991) also measured the higher biomass (4.18 g/m^2) of *Galium* spp. in the herbaceous vegetation type. The lower *Galium* spp. biomass or abundance was found in the bamboo. When it is not disturbed, the bamboo forest in the Virunga forms closed canopies with very little undergrowth (Vedder, 1984). The stepwise regression model showed an increase in *Galium* spp. abundance or biomass with increased altitude; but becomes lower with higher Hagenia tree densities. The *Galium* spp. BRT model (AUC=0.65) predicted the higher probabilities (>0.60) of occurrence of *Galium* spp. in the western facing slopes and high solar radiation (figure 20). The moist western-facing slopes of the Virunga support high vegetation abundance (Sternberg & Shoshany, 2001).

5.3.3. ***Carduus nyassanus*** **(thistle)**

The *C. nyassanus* biomass was significantly different between two vegetation groups. The higher biomass (> 5 g/m^2) of this species was found in the brush ridge, Hagenia-Hypericum and herbaceous vegetation zones. The higher biomass 25.40 g/m^2) of *C. nyassanus* was also measured by Plumptre (1991) in the brush ridge. The stepwise regression showed that *C. nyassanus* biomass significantly increased in higher slope values and declines in the eastern-facing slopes. The moist western-facing slopes of the Virunga support high vegetation abundance (Sternberg & Shoshany, 2001). The *C. nyassanus* BRT model (AUC=0.77) predicted the higher probabilities of occurrence of this species in the elevation range of 3000 m – 3400 m and on the west-facing slopes. These elevation zones have also been reported to be suitable for mountain gorilla habitat (van Gils & Kayijamahe, 2010).

5.3.4. ***Peucedanum linderi*** **(wild celery)**

The *P. linderi* biomass or abundance was not statistically significant between vegetation types (p> 0.05). This species was found to only occur in the Hagenia-Hypericum, herbaceous, mixed forest and brush ridge. Plumptre (1991) also measured higher biomass (24.37 g/m^2) of this species in the herbaceous vegetation. The stepwise regression showed that *P. linderi* biomass was significantly lower in higher elevations and increased in high values of both NDVI and EVI2. The *P. linderi* BRT model (AUC=0.72) predicted the higher probabilities of occurrence of this species in the elevations lower than 3300 m, western-facing slopes and open forest canopies (figure 19). Watts (1984) confirms that *P. linderi* was abundant in elevations below 3200 m. Dondeyne et al. (1993) confirms that due the very cold conditions on the volcano summit, the organic matter decomposes slowly and the available plant nutrients are scarce. This explains why the *P. linderi* does not occur in higher elevations of the Virunga massif.

5.3.5. ***Laportea alatipes*** **(nettles)**

The *L. alatipes* was biomass or abundance was not statistically significant between vegetation types (p>0.05). The higher biomass (> 5g /m^2) of this species was found was found in the Hagenia-Hypericum, brush ridge, herbaceous and bamboo. Plumptre (1991) also measured higher biomass (60.19g/m^2 and 57.59 g/m^2) of this species in the herbaceous and Hagenia-Hypericum vegetation. The stepwise regression showed a significant decrease in *L. alatipes* biomass with increased elevation and eastness (table 15). Watts (1984) also found the *L. alatipes* to be abundant in altitudes below 3200 m. The current study showed that this species biomass was also significantly increasing in higher NDVI and EVI2. The *L. alatipes* BRT model (AUC=0.80) predicted the higher probabilities of occurrence of this species in the elevation ranges of 2800 m – 3200 m, in Hagenia-Hypericum vegetation type and on the western-facing slopes (figure 19). The Hagenia-Hypericum vegetation type was also reported to predict the most suitable areas for mountain gorillas (van Gils & Kayijamahe, 2010).

6. CONCLUSION AND RECOMMENDATIONS

The main purpose of this study was to assess the power of remote sensing-derived variables to predict the abundance or biomass and spatial distribution of mountain gorilla food species in the Virunga massif. The continuous monitoring of these species is of great importance since the number of mountain gorillas has increased since 1981. There were two dataset: dataset 1 (abundance) collected by the researcher in October 2015 and dataset 2 (biomass) provided by Dr. Grueter Cyril in collaboration with the Dian Fossey Gorilla Fund International.

6.1. General conclusion

The Analysis of Variance (ANOVA) showed significant difference ($p < 0.05$) in three gorilla food species biomass between vegetation types. The stepwise multiple linear regression models were significant ($p < 0.05$) for all five species in dataset 2, whereas significant for only two species in dataset 1. The non-significant results were attributed to the small number of observations for each of the species in dataset 1.
Two gorilla food species abundance or biomass was not explained by the vegetation indices. The Boosted Regression Tree species distribution model performed with a reasonable AUC (> 0.70) for four species and hardly greater than random guess (AUC=0.65) for one species.

The Virunga forest canopy cover was predicted by the three Aster bands namely band 2, 3 and 7. The results of this study showed a significant negative correlation between the canopy cover and the Aster band 3; but with lower variance explained by the model ($R^2=0.128$). The Aster imagery poorly predicted the Virunga forest canopy cover.

6.2. Five mountain gorilla food species abundance or biomass and spatial distribution

Rubus spp. biomass was significantly different between vegetation types: alpine and meadow, brush ridge and Hagenia-Hypericum, Hagenia-Hypericum and sub-alpine, Hagenia-Hypericum and meadow, etc.
This species biomass and abundance was significantly increasing in higher elevations as well as high Hypericum tree densities. The *Rubus* spp. Boosted Regression Tree model (AUC=0.78) showed that the elevation, eastness and the slope are most important variables predicting the occurrence of this species. Its optimum occurrence is in elevation ranges of 3200 m – 3600 m and in the negative eastness (western-facing slopes).

Galium spp. biomass was statistically significant between vegetation types: brush ridge and meadow as well as meadow and sub-alpine. This species abundance and biomass was significantly increasing in higher elevations. The *Galium* spp. Boosted Regression Tree model (AUC=0.65) showed that eastness, elevation and solar radiation are the most important variables predicting the occurrence of *Galium* spp. in the Virunga massif. The AUC value closer to 0.5 suggests that the model could not easily distinguish the species presence and absence; which means that the species is found everywhere in the study area.

P. linderi biomass or abundance was not statistically significant between vegetation types. This species biomass was significantly ($p < 0.05$) decreasing in higher elevations. The Boosted Regression Tree model (AUC=0.72) showed that the eastness, elevation and forest canopy cover are the most important predictors of *P. linderi* occurrence. Its higher probability of occurrence was found in the elevation range below 3200 m, western-facing slopes and open forest canopies.

C. nyassanus biomass was statistically significant between vegetation types: brush ridge and herbaceous as well as herbaceous and sub-alpine. This species biomass was significantly ($p<0.05$) lower in the east-facing slopes but higher in high slope percentages. The Boosted Regression Tree model (AUC=0.77) showed that the elevation, eastness and northness are most relevant variables predicting the occurrence of *C. nyassanus* in the Virunga massif.

L. alatipes abundance or biomass was not statistically significant ($p >0.05$) between vegetation types. This species biomass was significantly ($p< 0.05$) decreasing with increased altitude and eastness. The Boosted Regression Tree model (AUC=0.80) showed that the elevation, vegetation types and eastness are the most important variables predicting the occurrence or *L. alatipes* in the Virunga massif.

6.3. Research limitations and recommendations

The majority of the species abundance/biomass and presence data used in this study was collected in the Karisoke area in the saddles between the Karisimbi and Bisoke volcanoes (Figure 9). Although this zone is already representative in terms of vegetation types; the gorilla food species could also be recorded and analysed in all zones of the Virunga massif. More emphasis should be on examining the quality of the *Rubus* spp. or *Galium* spp. species found in the sub-alpine and alpine zones.

The current study analysed two different datasets; the dataset 1 (collected only during three weeks) was small compared to the dataset 2 (collected the whole year). The smaller the dataset, the higher the probability of not getting a significant difference if one really exists. The dataset 2 was collected using the random sampling while the dataset 1 was collected using the systematic sampling. The latter requires enough time to collect at least 30 observations per species in each vegetation type. Large sample size collection are not easily achievable within a short period of time and in a rugged terrain such as the Virunga massif.

The satellite imagery used in the current study was not covering the eastern side of the Virunga massif, due to cloud problems. Although they are cost-intensive and require the expert for their processing and analysis; there are sensors (e.g. Radar) which are hardly affected by either clouds, dust, fog, wind or any bad weather conditions. Hence, the free-cloud imageries should be useful for the application of satellite imagery data to study the mountain gorilla habitat properties.

This study showed the spatial distribution maps of five mostly consumed mountain gorilla food species in the Virunga massif; but the stepwise variable reduction approach was not followed. The gorillas are supposed to be found where their food is abundant. Whilst the species distribution models based on climate change scenarios predict the future gorilla habitat suitability; there is need to compare the current and past time gorilla food species distribution. This will facilitate the park managers to know the movement of the mountain gorillas with regard to the changes in their food availability. It could be better to use the best predictor selection approach and model each species separately and without comparing species. Moreover, weather stations exist in the Karisoke area (Tuyisingize, 2010); instruments for measuring light intensity can be established. By including the soil, measured light intensity and grazing pressure parameters in the model; it can improve the accuracy.

LIST OF REFERENCES

Abrams, M., Tsu, H., Hulley, G., Iwao, K., Pieri, D., Cudahy, T., & Kargel, J. (2015). The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) after fifteen years: Review of global products. *International Journal of Applied Earth Observation and Geoinformation*, *38*, 292–301. doi:10.1016/j.jag.2015.01.013

Advani, N. (2014). *WWF Wildlife and Climate Change Series: Mountain gorilla. World Wildlife Fund.* Washington, D.C.

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*(6), 1223–1232. doi:10.1111/j.1365-2664.2006.01214.x

Austin, M. . (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, *157*(2-3), 101–118. doi:10.1016/S0304-3800(02)00205-3

Bausch, W. C. (1993). Soil background effects on reflectance-based crop coefficients for corn. *Remote Sensing of Environment*, *46*(2), 213–222. doi:10.1016/0034-4257(93)90096-G

Belfiore, N., Seimon, A., Picton, P., G., Basabose, A., Gray, M., Masinde, I., Elliott, J., Thorne, J., H., Seo, C., W., Muruthi, P. (2015). *The Implications of Global Climate Change for Mountain Gorilla Conservation in the Albertine Rift. A White Paper prepared by the African Wildlife Foundation, the International Gorilla Conservation Programme and EcoAdapt.*

Bochet, E., & García-Fayos, P. (2004). Factors controlling vegetation establishment and water erosion on motorway slopes in Valencia, Spain. *Restoration Ecology*, *12*(2), 166–174. doi:10.1111/j.1061-2971.2004.0325.x

De' Ath, G., & Fabricius, K. E. (2013). Classification and Regression Trees : A Powerful Yet Simple Technique for Ecological Data Analysis. *America*, *81*(11), 3178–3192.

De'ath, G. (2007). Boosted Trees for Ecological Modeling and Prediction. *Ecological Society of America*, *88*(1), 243–251.

De'Ath, G., and Fabricus, E. K. (2000). Classification and Regression Trees : A Powerful Yet Simple Technique for Ecological Data Analysis. *Ecological Society of America*, *81*(11), 3178–3192.

Dondeyne, S., Deckers, J. A., & Chapelle, J. (1993). The soils and vegetation of the Bisoke volcano (Rwanda): habitat of mountain gorillas. *Pedologie*, 301–322.

Doran, D. M., & Mcneilage, A. (1998). Gorilla ecology and behavior. *Evolutionary Anthropology: Issues, News and Reviews*, *6*(4), 120–131. doi:10.1002/(SICI)1520-6505(1998)6:4<120::AID-EVAN2>3.3.CO;2-L

Duque-Lazo, J., van Gils, H., Groen, T. A., & Navarro-Cerrillo, R. M. (2016). Transferability of species distribution models: The case of Phytophthora cinnamomi in Southwest Spain and Southwest Australia. *Ecological Modelling*, *320*, 62–70. doi:10.1016/j.ecolmodel.2015.09.019

Elith, J., & Leathwick, J. R. (2009). Species Distribution Models : Ecological Explanation and Prediction Across Space and Time. doi:10.1146/annurev.ecolsys.110308.120159

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802–813. doi:10.1111/j.1365-2656.2008.01390.x

Ellenberg, D., & Mueller-Dombois, D. (1974). Community Sampling: The Relevé Method. *Aims and Methods of Vegetation Ecology*. John Wiley and Sons: New York London Sydney Toronto.

Fatiha, B., Abdelkader, A., Latifa, H., & Mohamed, E. (2013). Spatio temporal analysis of vegetation by vegetation indices from multi-dates satellite images: Application to a semi arid area in ALGERIA. *Energy Procedia*, *36*, 667–675. doi:10.1016/j.egypro.2013.07.077

Field, A. (2009). *Discovering Statistics Using SPSS* (Third Edit.). London-Thousand Oaks-New Delhi-Singapore: SAGE Publications.

Fossey, D. (1974). Observations on the home range of one group of mountain gorillas (Gorilla gorilla beringei). *Animal Behaviour*, *22*(3), 568–581. doi:10.1016/S0003-3472(74)80002-3

Foster, P. (2001). The potential negative impacts of global climate change on tropical montane cloud forests. *Earth-Science Reviews*, *55*(1-2), 73–106. doi:10.1016/S0012-8252(01)00056-3

Fu, W. J., Jiang, P. K., Zhou, G. M., & Zhao, K. L. (2014). Using Moran's I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China. *Biogeosciences*, *11*(8), 2401–2409. doi:10.5194/bg-11-2401-2014

Ganas, J., Ortmann, S., & Robbins, M. M. (2008). Food preferences of wild mountain gorillas. *American Journal of Primatology*, *70*(10), 927–938. doi:10.1002/ajp.20584

Garrison, P. J., Marshall, D. W., Stremick-thompson, L., Cicero, P. L., & Dearlove, P. D. (2005). *Effects of Pier Shading on Littoral Zone Habitat and Communities in Lakes Ripley and Rock , Jefferson County ,. PUB-SS-1006 2005.*

Gebreslasie, M. T., Ahmed, F. B., & van Aardt, J. (2009). Image-based reflectance conversion of ASTER and IKONOS imagery as precursor to structural assessment of plantation forests in KwaZulu-Natal, South Africa. *Southern Forests: A Journal of Forest Science*, *71*(4), 259–265. doi:10.2989/SF.2009.71.4.2.1029

Ghulam, A. (2009). How to calculate reflectance and temperature using ASTER data. Center for Environmental Sciences at Saint Louis University.

Goward, S. N., Markham, B., Dye, D. G., Dulaney, W., & Yang, J. (1991). Normalized difference vegetation index measurements from the advanced very high resolution radiometer. *Remote Sensing of Environment*, *35*(2-3), 257–277. doi:10.1016/0034-4257(91)90017-Z

Gray, M., Roy, J., Vigilant, L., Fawcett, K., Basabose, A., Cranfield, M., … Robbins, M. M. (2013). Genetic census reveals increased but uneven growth of a critically endangered mountain gorilla population. *Biological Conservation*, *158*, 230–238. doi:10.1016/j.biocon.2012.09.018

Grueter, C. C., Ndamiyabo, F., Plumptre, A. J., Abavandimwe, D., Mundry, R., Fawcett, K. a., & Robbins, M. M. (2013). Long-Term Temporal and Spatial Dynamics of Food Availability for Endangered Mountain Gorillas in Volcanoes National Park, Rwanda. *American Journal of Primatology*, *75*(3), 267–280. doi:10.1002/ajp.22102

Gu, Y., Wylie, B. K., Howard, D. M., Phuyal, K. P., & Ji, L. (2013). NDVI saturation adjustment: A new approach for improving cropland performance estimates in the Greater Platte River Basin, USA. *Ecological Indicators*, *30*, 1–6. doi:10.1016/j.ecolind.2013.01.041

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*, 147–186. doi:10.1016/S0304-3800(00)00354-9

Hastie, T., Tibshirani, R. & Friedman, J. H. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer-Verlag, New York.

Hitimana, J., Namara, A., & Sengalama, T. (2006). *Community-Based Natural Resource Management (CBNRM) Plan. Report prepared for the International Gorilla Conservation Programme.*

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195–213. doi:10.1016/S0034-4257(02)00096-2

Huete, A. R., Jackson, R. D., & Post, D. F. (1985). Spectral response of a plant canopy with different soil backgrounds. *Remote Sensing of Environment*, *17*(1), 37–53. doi:10.1016/0034-4257(85)90111-7

Jiang, Z., Huete, A., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, *112*(10), 3833–3845. doi:10.1016/j.rse.2008.06.006

Joshi, C., De Leeuw, J., van Andel, J., Skidmore, A. K., Lekhak, H. D., van Duren, I. C., & Norbu, N. (2006). Indirect remote sensing of a cryptic forest understorey invasive species. *Forest Ecology and Management*, *225*(1-3), 245–256. doi:10.1016/j.foreco.2006.01.013

Joshi, C., De Leeuw, J., & van Duren, I. C. (2004). Remote Sensing and GIS Applications for Mapping Spatial Modelling of Invasive Spesies. In *Proceedings of the XXth ISPRS congress : Geo-imagery bridging* (Vol. 35, pp. 669–677). Istanbul, Turkey.

Joshi, C., Leeuw, J. De, Skidmore, A. K., van Duren, I. C., & van Oosten, H. (2006). Remotely sensed estimation of forest canopy density: A comparison of the performance of four methods. *International Journal of Applied Earth Observation and Geoinformation*, *8*(2), 84–95. doi:10.1016/j.jag.2005.08.004

Kayijamahe, E. (2008). *Spatial modelling of mountain gorilla (Gorilla beringei beringei) habitat suitability and human impact. MSc Thesis. University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC).* Retrieved from http://www.itc.nl/library/papers_2008/msc/nrm/kayijamahe.pdf

Kern, C. C., Montgomery, R. A., Reich, P. B., & Strong, T. F. (2013). Canopy gap size influences niche partitioning of the ground-layer plant community in a northern temperate forest. *Journal of Plant Ecology*, *6*(1), 101–112. doi:10.1093/jpe/rts016

Leathwick, J. R., Elith, J., Francis, M. P., Hastie, T., & Taylor, P. (2006). 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. doi:10.3354/meps321267

Legendre, P. (1993). Spatial Autocorrelation: Trouble or New Paradigm? *Ecological Society of America*, *74*(6), 1659–1673.

Luke, W., R., Q. (2010). Carduus nyassanus. The IUCN Red List of Threatened Species 2010: e.T185589A8442144. doi:http://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T185589A8442144.en

Måren, I. E., Karki, S., Prajapati, C., Yadav, R. K., & Shrestha, B. B. (2015). Facing north or south: Does slope aspect impact forest stand characteristics and soil properties in a semiarid trans-Himalayan valley? *Journal of Arid Environments*, *121*, 112–123. doi:10.1016/j.jaridenv.2015.06.004

Martens, S. N., Breshears, D. D., & Meyer, C. W. (2000). Spatial distributions of understory light along the grassland/forest continuum: Effects of cover, height, and spatial pattern of tree canopies. *Ecological Modelling*, *126*(1), 79–93. doi:10.1016/S0304-3800(99)00188-X

MINECOFIN. (2012). *Rwanda fourth population and housing census. Population size, structure and distribution*. Kigali,Rwanda.

Naumburg, E., & Dewald, L. E. (1999). Relationships between Pinus ponderosa forest structure , light characteristics , and understory graminoid species presence and abundance. *Forest Ecology and Management*, *124*, 205–215.

Nyandwi, E. (2008). *Road edge effect on Forest Canopy structure and Epiphyte biodiversity in a Tropical Mountainous Rainforest, Nyungwe National Park, Rwanda, MSc Thesis*. University of Twente, Faculty of Geo-information science and Earth Observation.

Owiunji, I., Nkuutu, D., Kujirankwinja, D., Liengola, I., Plumptre, A., Nsanzurwimo, A., … McNeilage, A. (2005). *The biodiversity of the Virunga Volcanoes*.

Piedallu, C., & Gégout, J. (2008). Efficient assessment of topographic solar radiation to improve plant distribution models. *Agricultural and Forest Meteorology*, *148*(11), 1696–1706. doi:10.1016/j.agrformet.2008.06.001

Pinkerton, M. H., Smith, A. N. H., Raymond, B., Hosie, G. W., Sharp, B., Leathwick, J. R., & Bradford-Grieve, J. M. (2010). Spatial and seasonal distribution of adult Oithona similis in the Southern Ocean: Predictions using boosted regression trees. *Deep-Sea Research Part I: Oceanographic Research Papers*, *57*(4), 469–485. doi:10.1016/j.dsr.2009.12.010

Plumptre, A. (1991). *Plant-Herbivore dynamics in the Birungas. PhD Thesis*. University of Bristol.

Plumptre, A., Davenport, T., Behangana, M., Kityo, R., Eilu, G., Ssegawa, P., … Herremans, M. (2007). The biodiversity of the Albertine Rift. *Biological Conservation*, *134*(2), 178–194. doi:10.1016/j.biocon.2006.08.021

Plumptre, A. J. (1996). Modelling the impact of large herbivores on the food supply of mountain gorillas and implications for management. *Biological Conservation*, *75*(2), 147–155. doi:http://dx.doi.org/10.1016/0006-3207(95)00061-5

Plumptre, A., Kujirakwinja, D., Treves, A., Owiunji, I., & Rainer, H. (2007). Transboundary conservation in the greater Virunga landscape: Its importance for landscape species. *Biological Conservation*, *134*(2), 279–287. doi:10.1016/j.biocon.2006.08.012

Quinn, G. P., & Keough, M. J. (2002). *Experimental Design and Data Analysis for Biologists*. New York, USA: Cambridge University Press. doi:10.1016/S0022-0981(02)00278-2

Raevel, V., Munoz, F., Pons, V., Renaux, A., Martin, A., & Thompson, J. D. (2013). Changing assembly processes during a primary succession of plant communities on Mediterranean roadcuts. *Journal of Plant Ecology*, *6*(1), 19–28. doi:10.1093/jpe/rts011

REMA. (2011). *Atlas of Rwanda's Changing Environment: Implications for climate change resilience*. Kigali. Retrieved from https://na.unep.net/siouxfalls/publications/REMA.pdf

Richards, J. A., & Jia, X. (2006). *Remote Sensing Digital Image Analysis* (Fourth Edi.). Germany: Springer Berlin Heidelberg New York.

Robbins, M. M., Sicotte, P., & Stewart, K. J. (2005). *Mountain Gorillas: Three Decades of Research at Karisoke*. Cambridge University Press.

Rogan, J., Chen, D. (2004). Remote sensing technology for mapping and monitoring land-cover and land-use. *Progress in Planning*, *61*(4), 301–25. doi:10.1016/S0305-9006(03)00066-7

Smith, J.,B., M. (1977). Vegetation and Microclimate of East- and West-Facing Slopes in the Grasslands of MT Wilhelm, Papua New Guinea. *Journal of Ecology*, *65*(1), 39–53.

Spinage, C. A. (1972). The ecology and problems of the Volcano National Park, Rwanda. *Biological Conservation*, *4*(3), 194–204. doi:10.1016/0006-3207(72)90169-3

Sternberg, M., & Shoshany, M. (2001). Influence of slope aspect on Mediterranean woody formations: Comparison of a semiarid and an arid site in Israel. *Ecological Research*, *16*, 335–345. doi:10.1046/j.1440-1703.2001.00393.x

Tang, L., Wan, K., Cheng, C., Li, R., Wang, D., Pan, J., … Chen, F. (2014). Effect of fertilization patterns on the assemblage of weed communities in an upland winter wheat field. *Journal of Plant Ecology*, *7*(1), 39–50. doi:10.1093/jpe/rtt018

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. doi:10.1016/0034-4257(79)90013-0

Tuyisingize, D. (2010). *Terrestrial small mammal community composition in the Volcanoes National Park, Rwanda. MSc Thesis*. University of Cape Town.

van Gils, H., & Kayijamahe, E. (2010). Sharing natural resources: Mountain gorillas and people in the Parc National des Volcans, Rwanda. *African Journal of Ecology*, *48*(3), 621–627. doi:10.1111/j.1365-2028.2009.01154.x van Gils, H., Westinga, E., Carafa, M., Antonucci, A., & Ciaschetti, G. (2014). Where the bears roam in Majella National Park, Italy. *Journal for Nature Conservation*, *22*(1), 23–34. doi:10.1016/j.jnc.2013.08.001

van Gils, H., Conti, F., Ciaschetti, G., Westinga, E. (2012). Fine resolution distribution modelling of endemica in Majella National Park, Central Italy. *Plant Biosystems-An International Journal Dealing with All Aspects of Plant Biology: Official Journal of the Societa Botanica Italiana*, (146:sup1), 276–287. doi:10.1080/11263504.2014.976290

Vedder, A. L. (1984). Movement patterns of a group of free-ranging mountain gorillas (Gorilla gorilla beringei) and their relation to food availability. *American Journal of Primatology*, *7*(2), 73–88. doi:10.1002/ajp.1350070202

Veloz, S. D. (2009). Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, *36*(12), 2290–2299. doi:10.1111/j.1365-2699.2009.02174.x

Wang, T. J. (2009). *Observing giant panda habitat and forage abundance from space. PhD Thesis*. Faculty of Geo-inofrmation Science and Earth Observation, University of Twente.

Watts, D. P. (1984). Composition and variability of mountain gorilla diets in the Central Virungas. *American Journal of Primatology*, *7*(4), 323–356. doi:10.1002/ajp.1350070403

Watts, D. P. (1985). Relations between group size and composition and feeding competition in mountain gorilla groups. *Animal Behaviour*, *33*(1), 72–85. doi:10.1016/S0003-3472(85)80121-4

Watts, D. P. (1998a). Long-term habitat use by mountain gorillas (Gorilla gorilla beringei). 1. Consistencey, variation, and home range size and stability. *International Journal of Primatology*, *19*(4), 651–679.

Watts, D. P. (1998b). Long-term habitat use by mountain gorillas (Gorilla gorilla beringei). 2. Reuse of Foraging Areas in Relation to Resource Abundance, Quality and Depletion. *International Journal of Primatology*, *19*(4), 651–679. doi:10.1023/A:1020324909101

Watts, D. P. (1998c). Seasonality in the Ecology and Life Histories of Mountain Gorillas (Gorilla gorilla beringei). *International Journal of Primatology*, *19*(6), 929–948.

Weber, A. W., & Vedder, A. (1983). Population dynamics of the virunga gorillas: 1959–1978. *Biological Conservation*, *26*(4), 341–366. doi:10.1016/0006-3207(83)90096-4

Wikum, D. A., & Shanholtzer, G. F. (1978). Application of Braun-Blaquet cover-abundance scale for vegetation analysis in land-development studies. *Environmental Management*, *2*(4), 323–329. doi:10.1007/bf01866672

Yamaguchi, Y., Kahle, A. B., Tsu, H., Kawakami, T., & Pniel, M. (1998). Overview of advanced spaceborne thermal emission and reflection radiometer (ASTER). *IEEE Transactions on Geoscience and Remote Sensing*, *36*(4), 1062–1071. doi:10.1109/36.700991

APPENDICES

Appendix 1: Slope correction table

Slope correction table

Plot size 500 m^2

Slope%	Radius(m)	Slope%	Radius(m)	Slope%	Radius(m)
0	12.62				
1	12.62	36	13.01	71	13.97
2	12.62	37	13.03	72	14.00
3	12.62	38	13.05	73	14.04
4	12.62	39	13.07	74	14.07
5	12.62	40	13.09	75	14.10
6	12.63	41	13.12	76	14.14
7	12.63	42	13.14	77	14.17
8	12.64	43	13.16	78	14.21
9	12.64	44	13.19	79	14.24
10	12.65	45	13.21	80	14.28
11	12.65	46	13.24	81	14.31
12	12.66	47	13.26	82	14.35
13	12.67	48	13.29	83	14.38
14	12.68	49	13.31	84	14.42
15	12.69	50	13.34	85	14.45
16	12.70	51	13.37	86	14.49
17	12.71	52	13.39	87	14.52
18	12.72	53	13.42	88	14.56
19	12.73	54	13.45	89	14.60
20	12.74	55	13.48	90	14.63
21	12.75	56	13.51	91	14.67
22	12.77	57	13.53	92	14.71
23	12.78	58	13.56	93	14.74
24	12.79	59	13.59	94	14.78
25	12.81	60	13.62	95	14.82
26	12.82	61	13.65	96	14.85
27	12.84	62	13.68	97	14.89
28	12.86	63	13.72	98	14.93
29	12.87	64	13.75	99	14.97
30	12.89	65	13.78	100	15.00
31	12.91	66	13.81	101	15.04
32	12.93	67	13.84	102	15.08
33	12.95	68	13.87	103	15.12
34	12.97	69	13.91	104	15.15
35	12.99	70	13.94	105	15.19

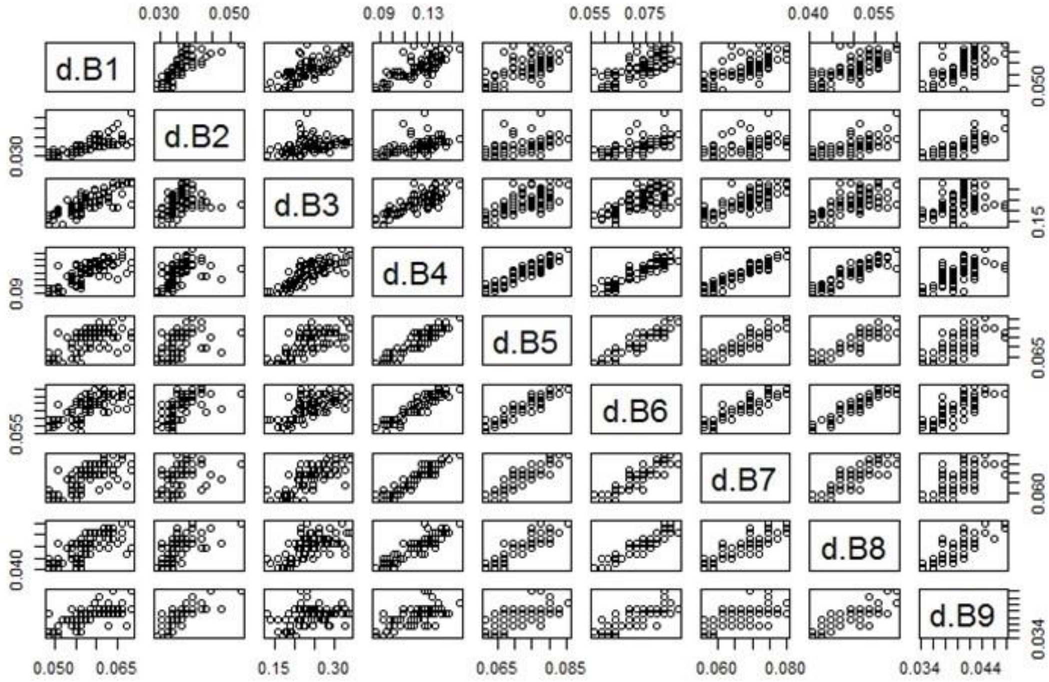
A. de Gier – 2000

Appendix 2: Field data collection sheet

<table>
<tr><td colspan="11">Vegetation data collection sheet in Volcanoes National Park</td><td>Sample No:</td></tr>
<tr><td colspan="2" rowspan="2">Date:</td><td colspan="2" rowspan="2">GPS

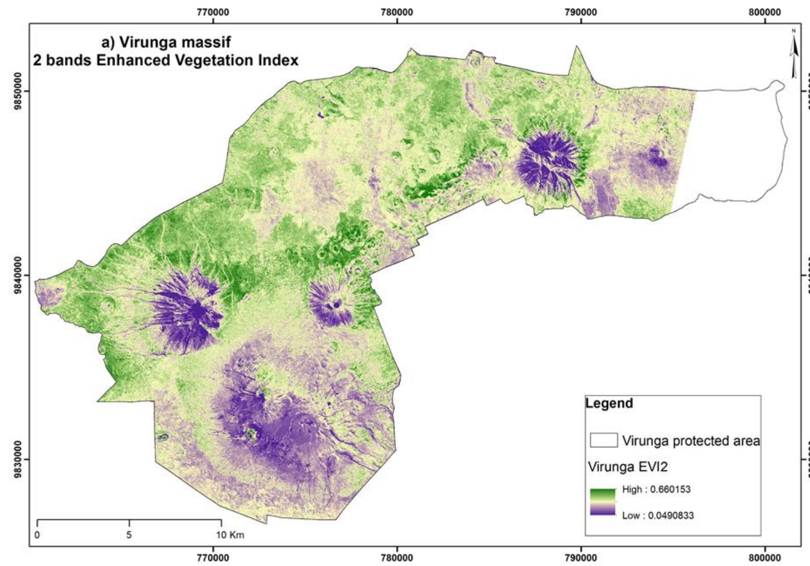
Coordinates</td><td>X</td><td></td><td></td><td></td><td></td><td></td><td></td><td>Observer name:</td></tr>
<tr><td>Y</td><td></td><td></td><td></td><td></td><td></td><td></td><td>Individual cover (%)</td></tr>
<tr><td>Plot size</td><td>dbh (m)</td><td>Canopy cover (%)</td><td>Tree height (m)</td><td>Forage abundance</td><td colspan="6">Dominant species</td><td></td></tr>
<tr><td rowspan="3">500 m^2 (trees)</td><td rowspan="3"></td><td rowspan="3">N:
S:
E:
W:</td><td rowspan="3"></td><td rowspan="3"></td><td colspan="6"></td><td></td></tr>
<tr><td colspan="6"></td><td></td></tr>
<tr><td colspan="6"></td><td></td></tr>
<tr><td rowspan="3">5 m^2 (shrubs)</td><td rowspan="3"></td><td rowspan="3"></td><td rowspan="3"></td><td rowspan="3"></td><td colspan="6"></td><td></td></tr>
<tr><td colspan="6"></td><td></td></tr>
<tr><td colspan="6"></td><td></td></tr>
<tr><td rowspan="3">1 m^2 (forage)</td><td rowspan="3"></td><td rowspan="3"></td><td rowspan="3"></td><td rowspan="3">N:
S:
E:
W:</td><td colspan="6"></td><td></td></tr>
<tr><td colspan="6"></td><td></td></tr>
<tr><td colspan="6"></td><td></td></tr>
<tr><td colspan="12">Other observations:
Stem density:</td></tr>
</table>

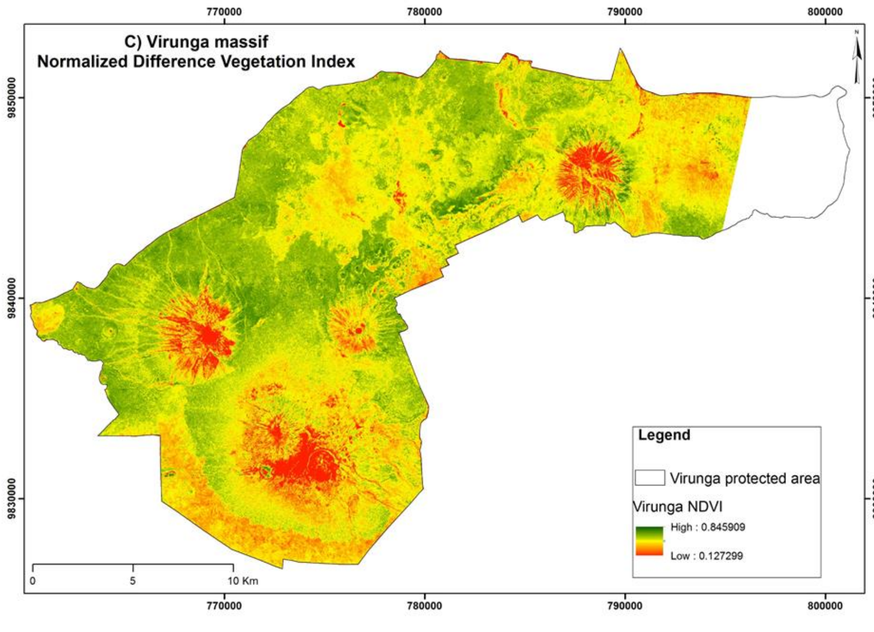
Appendix 3: Scatterplots showing the correlation between pairs of Aster bands reflectance values



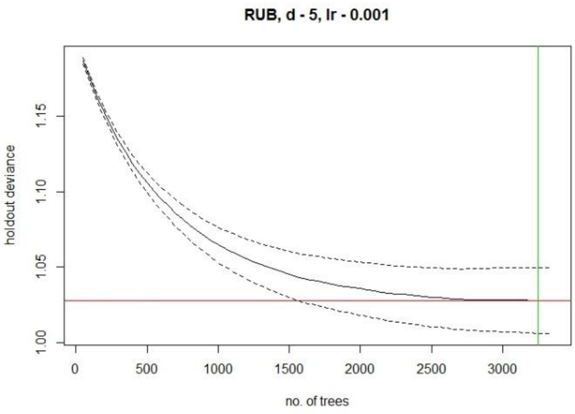
Appendix 4: Virunga vegetation indices maps[a].

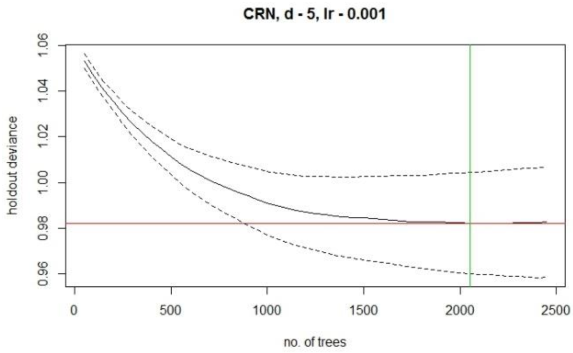
[a]The maps a and b are rotated 90o left.

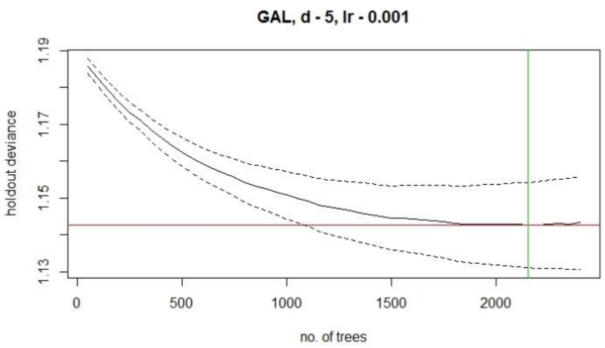


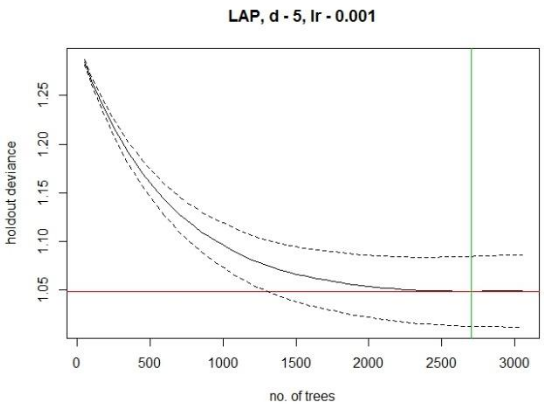


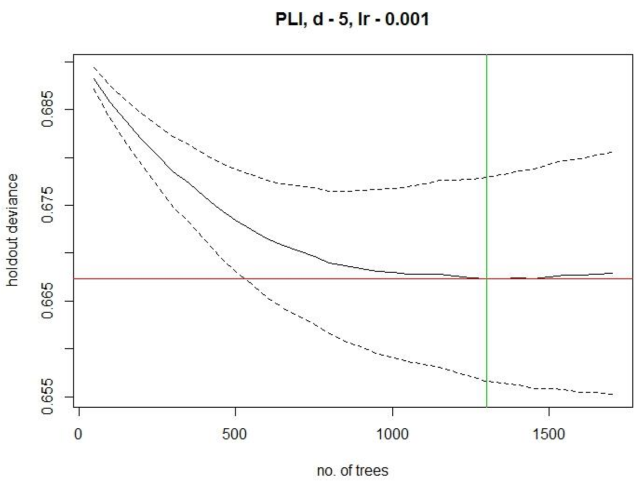
Appendix 5: *Rubus* spp., *C. nyassanus*, *Galium* spp., *L. alatipes* and *P. linderi* BRT model trees





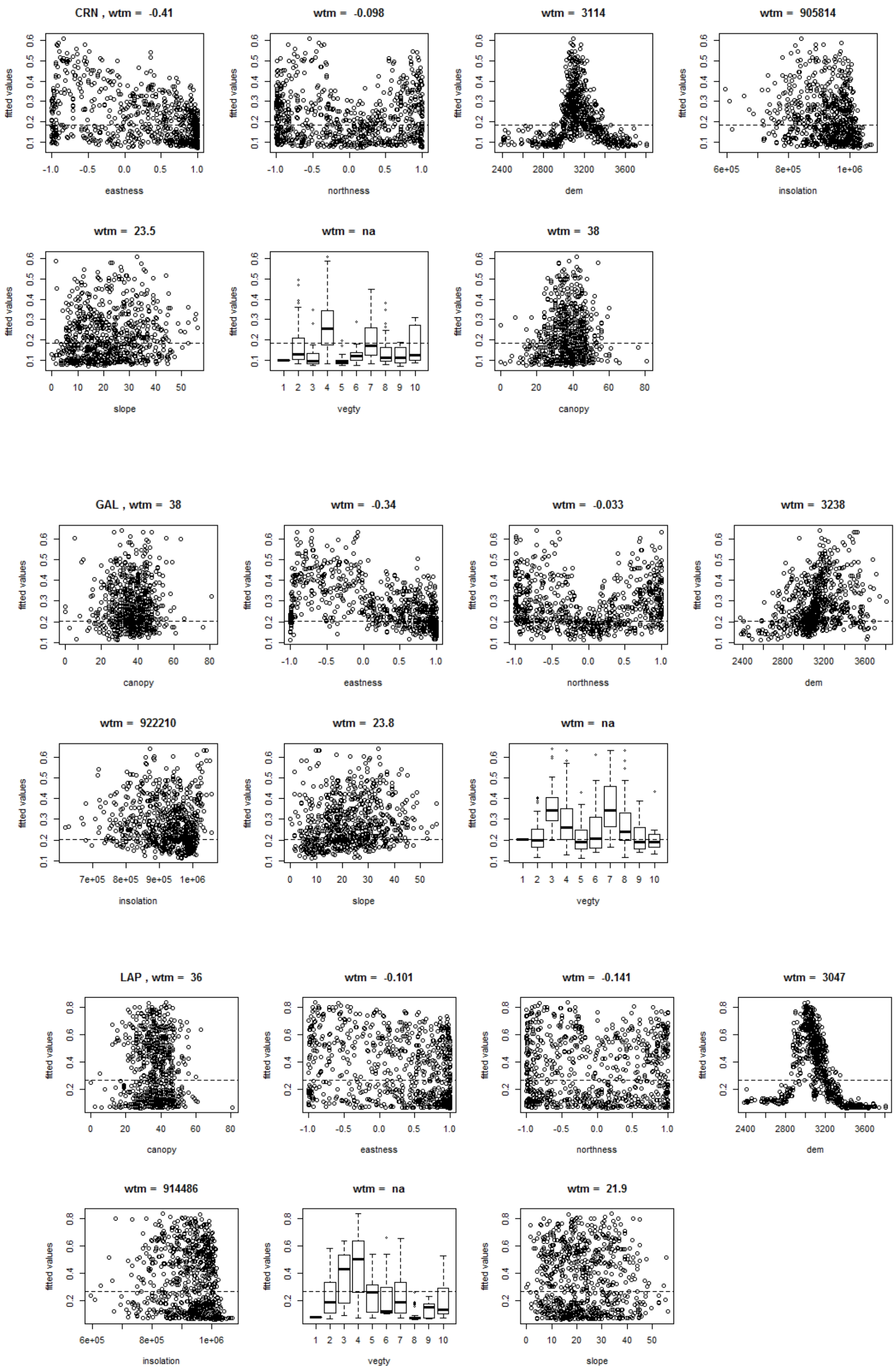






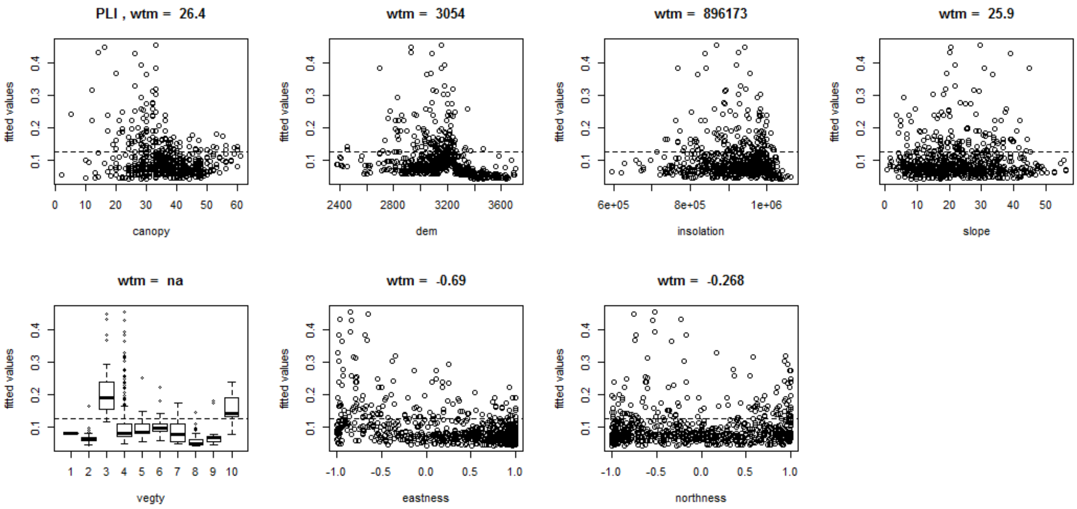
Appendix 6: Fitted values in relation to each predictor (three species)

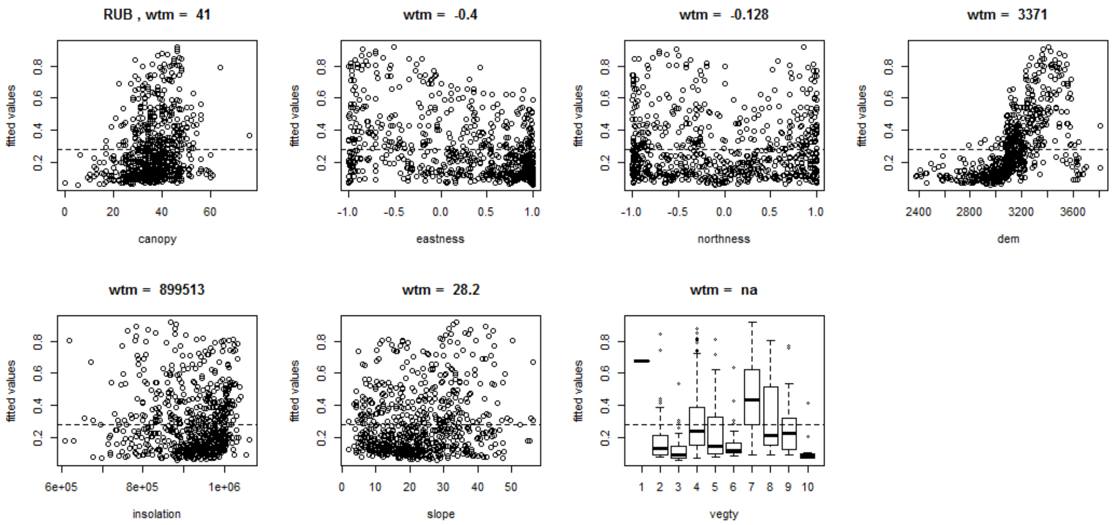
wtm: weighted mean of fitted values in relation to non-factor predictor. CRN: *C. nyassanus*; GAL: *Galium* spp.; LAP: *L. alatipes*. . For land cover types (vegty), 1: water; 2: bamboo; 3: herbaceous; 4: Hagenia-Hypericum; 5: Mixed forest; 6: Neoboutonia; 7: Brush ridge; 8: sub-alpine & alpine; 9: meadow/savannah; 10: Mimulopsis



Appendix 7: Fitted values in relation to each predictor (two species)

wtm: weighted mean of fitted values in relation to non-factor predictor. PLI: *P. linderi*; RUB: *Rubus* spp. For land cover types (vegty), 1: water; 2: bamboo; 3: herbaceous; 4: Hagenia-Hypericum; 5: Mixed forest; 6: Neoboutonia; 7: Brush ridge; 8: sub-alpine & alpine; 9: meadow/savannah; 10: Mimulopsis





Appendix 8: Pearson correlation matrices for both dataset 1 and 2[a].

[a]The first table (right side) with HAG & HYR biomass is for dataset 2 while the dataset 1 (left side) table contains HT, STD, DBH and CC standing for height, stem density, diameter and canopy cover respectively.

	HYR	HAG	slope	solar rad.	elevation	eastness	northness
HYR	1						
HAG	-0.017	1					
slope	-0.075	0.01	1				
solar rad.	0.05	-0.01	-0.92	1			
elevation	0.02	0.04	0.19	-0.001	1		
eastness	-0.03	0.02	0.11	-0.02	-0.008	1	
northness	0.001	-0.02	-0.13	0.21	-0.04	0.008	1

	slope	solar rad.	elevation	eastness	northness	HT	DBH	STD	CC
slope	1								
solar rad.	-0.75	1							
elevation	0.25	0.38	1						
eastness	0.04	-0.17	-0.28	1					
northness	0.04	0.20	0.27	0.20	1				
HT	-0.05	-0.01	-0.16	0.005	0.032	1			
DBH	-0.01	-0.02	-0.16	0.17	0.14	0.77	1		
STD	-0.17	-0.01	-0.27	-0.01	-0.03	0.43	0.15	1	
CC	-0.17	-0.16	-0.49	0.10	-0.16	0.53	0.39	0.43	1

Appendix 9: Pearson correlation coefficients between pairs of Aster band reflectance values.

	B1	B2	B3	B4	B5	B6	B7	B8	B9
B1	1								
B2	0.75	1							
B3	0.74	0.30	1						
B4	0.74	0.48	0.79	1					
B5	0.63	0.46	0.68	0.93	1				
B6	0.69	0.59	0.65	0.92	0.93	1			
B7	0.69	0.46	0.74	0.94	0.92	0.90	1		
B8	0.70	0.62	0.62	0.90	0.89	0.91	0.89	1	
B9	0.67	0.74	0.37	0.67	0.65	0.73	0.66	0.79	1

Appendix 10: Confusion matrix showing DIC results (Aster 2006)[a].

[a]**Wt**: water; **Bab**: Bamboo; **Ha**: Hagenia-Hypericum; **Mf**: Mixed forest; **Neo**: Neoboutonia; **Br**: Brush ridge; **Ap**: Alpine; **Mim**: Mimulopsis; **Sav**: Savannah/Meadow.

Reference points

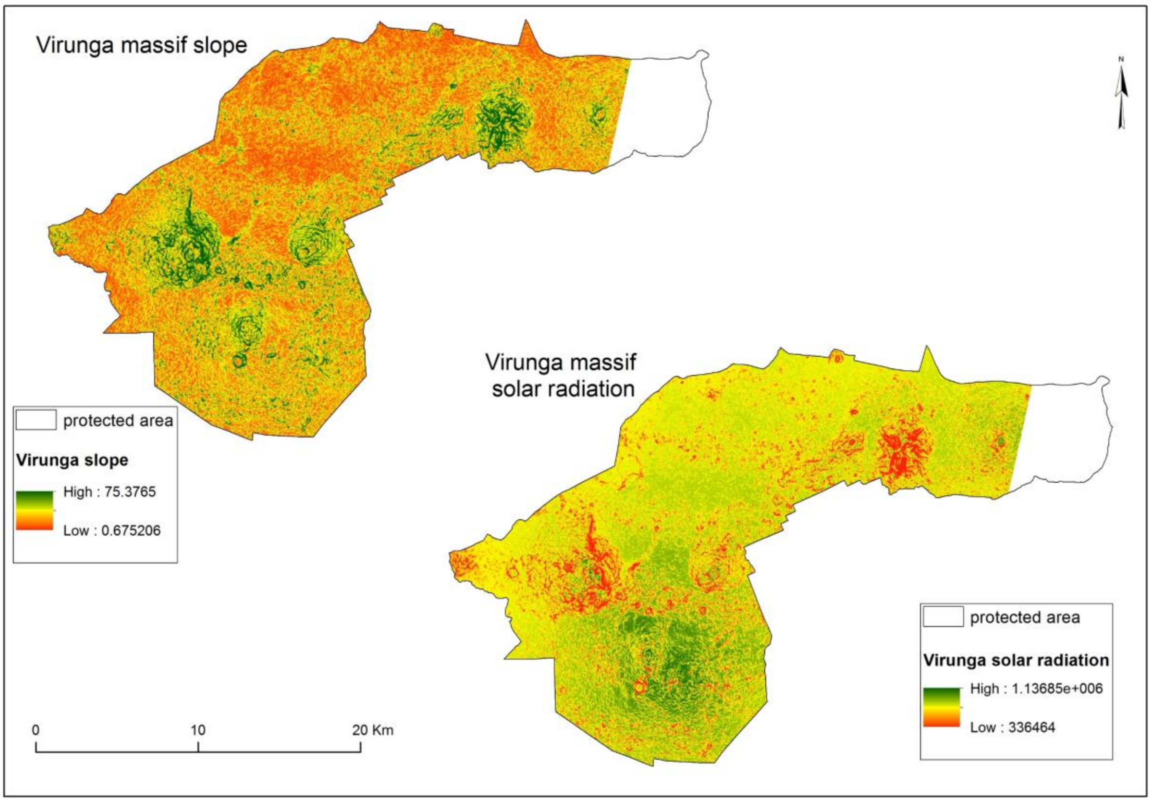
Classification results

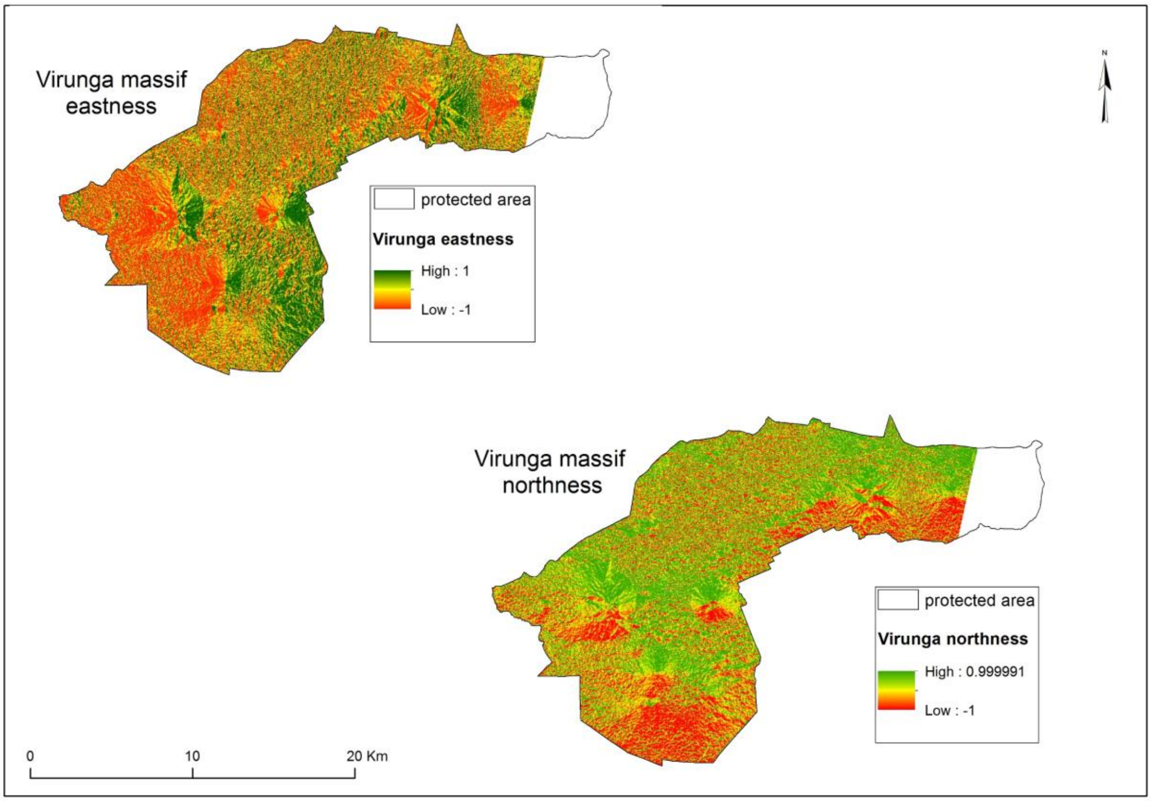
	Wt	Bab	Herb	Ha	Mf	Neo	Br	Ap	Me	Mim	Sav	Total
Wt	**1**	0	0	0	0	0	0	0	0	0	0	1
Bab	0	**26**	0	9	1	0	1	0	0	0	0	37
Herb	0	0	**7**	4	0	0	1	0	0	0	0	12
Ha	0	7	1	**159**	2	2	6	7	3	0	0	187
Mf	0	0	0	1	**10**	0	0	1	0	0	0	12
Neo	0	0	0	1	0	**5**	1	0	0	0	0	7
Br	0	1	0	3	0	0	**14**	4	0	0	0	22
Ap	0	0	0	2	0	0	0	**19**	1	0	0	22
Me	0	0	0	2	0	0	1	1	**12**	0	0	16
Mim	0	0	0	0	0	0	0	0	0	**2**	0	2
Total	1	34	8	181	13	7	24	32	10	2	6	318

Appendix 11: Secondary herbaceous formations with ferns dominant[a].

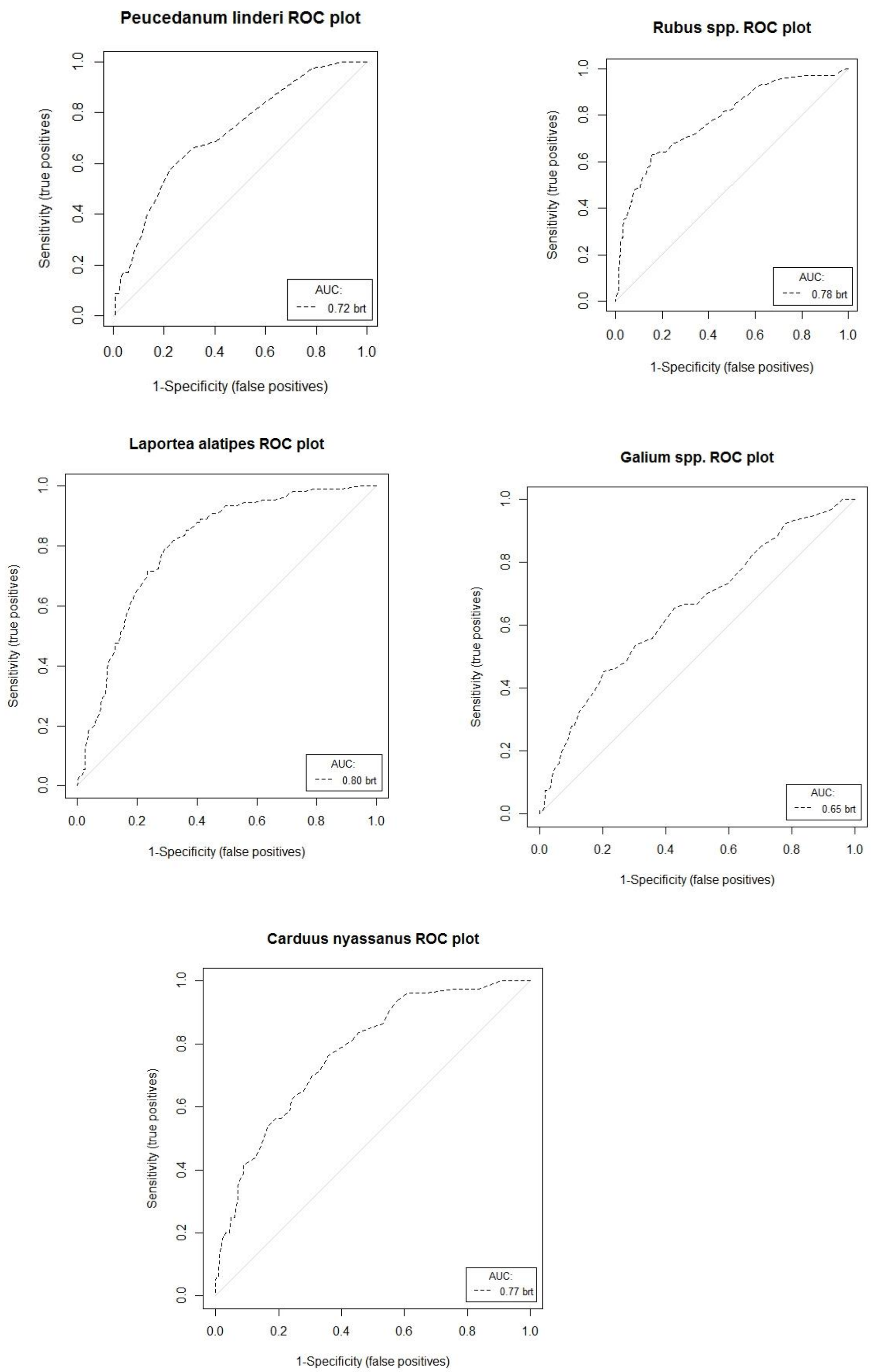
[a]Photo taken in Volcanoes National Park during fieldwork (October 2015).

Appendix 12: Virunga topography

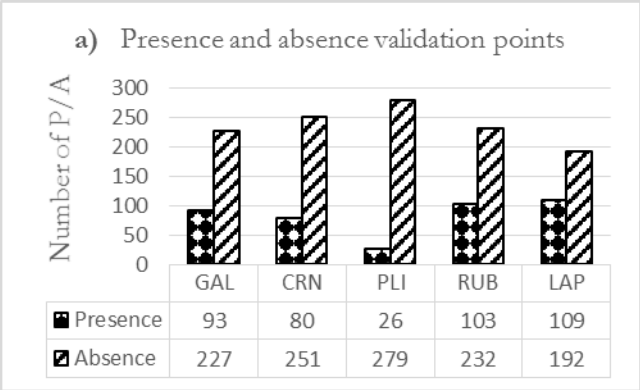




Appendix 13: ROC plots for the five gorilla food species BRT models



Appendix 14: Number of presence and absence used in the validation dataset



	GAL	CRN	PLI	RUB	LAP
Presence	93	80	26	103	109
Absence	227	251	279	232	192

Appendix 15: Spectral profile of different vegetation types and water in the Virunga protected area.

Nine bands of the Aster imagery where the pixel value reaches its maximum peak in band 3 (NIR)