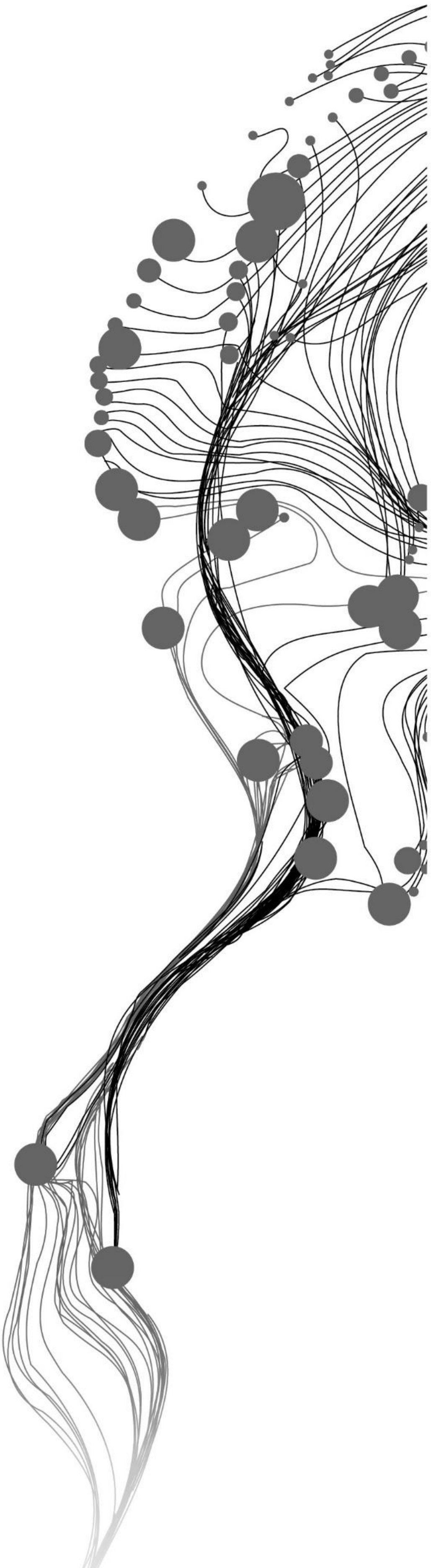
RISK ASSESSMENT OF BARK BEETLE OUTBREAK IN THE SCHWARZWALD NATIONAL PARK

FERNANDO FERNANDEZ PEREZ
June, 2020

SUPERVISORS:
Dr. I.C. van Duren
Drs. J.M. Looijen



RISK ASSESSMENT OF BARK BEETLE OUTBREAK IN THE SCHWARZWALD NATIONAL PARK

FERNANDO FERNANDEZ PEREZ
Enschede, The Netherlands, June, 2020

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 Spatial Engineering.

SUPERVISORS:
Dr. I.C. van Duren
Drs. J.M. Looijen

THESIS ASSESSMENT BOARD:
Dr. R. Darvish (Chair)
Prof.dr.ir. G.J.M.M. Nabuurs (External Examiner, Wageningen University)

ABSTRACT

The main threat of Norway spruces (Picea abies L.) in Europe is the bark beetle (Ips typographus L.). This insect is able to colonize weakened trees, but when there is a mass-attack, it can also breed in healthy trees. Forest management in Europe tries to minimize the severe effects of this pest. By knowing in which areas the outbreak probability is higher, the monitoring resources can be optimized. Modeling the distribution of bark beetles has been done before, obtaining different results in each study area. Therefore, our objective is to identify which spatial variables can be used for predicting the bark beetle outbreak in the Schwarzwald National Park. To that aim, a model created with boosted regression trees was applied. In order to avoid spatial autocorrelation, ten sub-models were calculated. The most important variables for predicting bark beetle outbreak in the study area were number of Norway spruces, height of Norway spruces, altitude, soil depth, slope, and percentage of Norway spruces compared with other tree species. These variables, in combination with other predictors, were used for creating the outbreak probability map. By using AUC as a validation method, the accuracy was 0.8. The model performance was also assessed with TSS, obtaining an accuracy of 0.49. This research provides insights into the spatial variables that can predict bark beetle outbreaks, which will support the decision-making process carried out by the National Park committee.

Keywords: Bark beetle outbreak, Ips typographus, BRT, SDM, Norway spruce, spatial variables

ACKNOWLEDGEMENTS

I would like to express my gratitude to dr. Iris van Duren and drs. Joan Looijen for giving me the opportunity of choosing this topic. I have learnt a lot during this process with their guidance, especially about critical thinking and scientific writing. Thanks for the interesting discussions and the kind words every time I got feedback. It has been a pleasure to be supervised by you.

I am very thankful to Dr. Markus Kautz for the interest in my research since the first moment, the help for obtaining the data, and his advice.

I also would like to thank dr.ir Thomas Groen, for the support with the model, R, and the constructive feedbacks and ideas.

My sincere thanks to dr. Christoph Dreiser from the Schwarzwald National Park for being open to share the data and allow this research.

Thanks go to my wonderful classmates from Spatial Engineering for these good moments in the ITC and for becoming such good friends. This journey has been a lot of fun with you.

Many thanks to Davina for your enormous support in my life and for pushing me when needed.

I would like to express my gratitude to Carina and Frank for treating me like one of the family and making me feel like home.

Finally, I want to thank my parents and my lovely sister for their encouragement and support in my life. They have always been a great role model and a source of motivation for me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Bark beetles (*Ips typographus* L.) outbreaks are the biggest biotic threat to Norway spruce (*Picea abies* L.) in Europe (Caudullo, Tinner, & Rigo, 2016). This pest can cause significant economic losses and reduce other ecosystem services provided by the forest (Hlásny, Krokene, & Liebhold, 2019). It is expected that with climate change, the spread of bark beetles will rise (Bentz & Jönsson, 2015; Ogris & Jurc, 2010; Seidl, Schelhaas, Lindner, & Lexer, 2009). Norway spruces are sensitive to this type of disturbance when they are weakened. Importantly, when the bark beetle population has enough breeding material in the forest, a mass outbreak can take place, also affecting the healthy trees (Caudullo et al., 2016; Christiansen & Bakke, 1988).

However, natural disturbances of the forest system are needed, since they are the drivers of the current forest ecosystems (Peltzer, Bast, Wilson, & Gerry, 2000). By means of disturbances, the abundance of forest species, the succession, and biodiversity are affected in the communities (Peltzer et al., 2000; Raffa et al., 2009).

1.1. Bark beetles ecology

In nature, there are around 6.000 different species of bark beetles (Hlásny et al., 2019). The most harmful species in Europe is the *Ips typographus L.* (Caudullo et al., 2016; Overbeck & Schmidt, 2012; Seidl, Schelhaas, & Lexer, 2011), which breeds in Norway spruces. Therefore, the bark beetle species of the *Ips typographus* will be the focus of the present study.

Bark beetle is considered a secondary pest, which means that it breeds in weakened trees (Lausch, Fahse, & Heurich, 2011). Hence, severe climatic events can damage trees, providing more breeding material, which leads to outbreaks of this insect (Giunta, Jenkins, Hebertson, & Munson, 2016).

Usually, Norway spruces have defenses against the colonization of the bark beetle, such as chemical, anatomical, and physiological (Hlásny et al., 2019). Christiansen & Bakke (1988) give the example of the oleoresin that is produced by the host tree. This oleoresin is expelled to sanitize and seal the wounds. The tree has the ability to defend itself against a certain amount of attacks. However, when a mass-attack takes place, the tree has not enough resources for producing defenses, and due to its amount, the bark beetle population is able to colonize healthy trees, as well (Christiansen & Bakke, 1988).

The Norway spruce represents not the only interaction that bark beetles have with other species. They are also associated with bluestain fungi, which uses the bark beetle for transport purposes. These microorganisms penetrate in the xylem of the tree, affecting the water flow (Hall, Castilla, White, Cooke, & Skakun, 2016). What the bark beetle gets in return is a source of nutrients for the larvae and protection against pathogens (Hlásny et al., 2019).

Natural enemies of bark beetles are predators (woodpeckers, flies, mites, mice, shrews, ants, and wasps), parasites (wasps, and nematodes), and pathogens (braconids and chalcids) (Christiansen & Bakke, 1988; Hlásny et al., 2019; Wegensteiner, Wermelinger, & Herrmann, 2015). As mentioned by Wegensteiner et al. (2015), these enemies hardly succeed in controlling the bark beetle population.

1.2. Population dynamics

The bark beetle attack is initiated in the weakened trees, in which the bark beetle is able to overcome the defenses of the trees (Hlásny et al., 2019). Nevertheless, as the population grows, the population attacks healthy trees and the symptoms of the attack can be classified into three different phases:

- Green attack: In this first phase, the beetle enters the inner bark and attracts the mates. The females create the galleries for laying the eggs, and eggs are laid on the bark of the host tree (Abdullah, 2019; Hlásny et al., 2019). In this phase, the trees do not show yet any visual sign of being infested on the foliage (Hlásny et al., 2019). Boreholes, sawdust, and loss of bark on the trunk can be observed (Fassnacht, Latifi, Ghosh, Joshi, & Koch, 2014). In addition, the bark beetle inoculates the blue fungi, which disrupts the water and nutrient flow through the xylem and phloem by means of its spores. This disruption affects the cooling process of the host, increasing, consequently, its temperature (Abdullah, 2019).
- Red attack: In this stage, the larvae hatch and feed on the bark by making tunnels into it. The larvae become adults and emerge from one tree to the ones, which are not with a high grade of infestation (Hlásny et al., 2019). The needles fade its color to yellow and then to red-brown (Abdullah, 2019; Fassnacht et al., 2014). This can last 1 or 2 years after the colonization of the tree (Hlásny et al., 2019).
- Grey attack: The effects on water and nutrient transport are critical, and the tree is not able to survive. The needles are lost (Hlásny et al., 2019).

The life cycle of bark beetles can be univoltine, which means that the generation is completed annually. In this case, the adults emerge from the tree and disperse to hibernation sites (Hlásny et al., 2019). Nonetheless, in areas of Central Europe, bivoltine populations can be found, which means that there is a second generation in the same year (Christiansen & Bakke, 1988). Within the second generation, populations are build-up faster (Hlásny et al., 2019). This high development is aided by the increase in temperature (Hlásny et al., 2019).

Two phases of a population can be differentiated (**Figure 1**). In the epidemic phase, the bark beetles are able to infest healthy trees. This is not the case in the endemic phase, in which most of the trees have enough defenses against the colonization, and the bark beetle population faces more difficulties (Hlásny et al., 2019). Therefore, at an endemic level, the colonization is influenced by the availability of hosts. In normal - or natural – conditions, the population is limited to endemic levels. Once the population is in an epidemic phase, only the lack of resources, the outbreak of natural enemies, or extreme temperatures can shift back the population to the endemic phase (Raffa, Grégoire, & Lindgren, 2015). The dynamics between both phases of the population (i.e., endemic and epidemic) change according to some external factors (e.g., windthrow, drought, or natural enemies). Importantly, most outbreaks are caused by a combination of different factors (Raffa et al., 2015), which can be classified as biotic and abiotic factors.

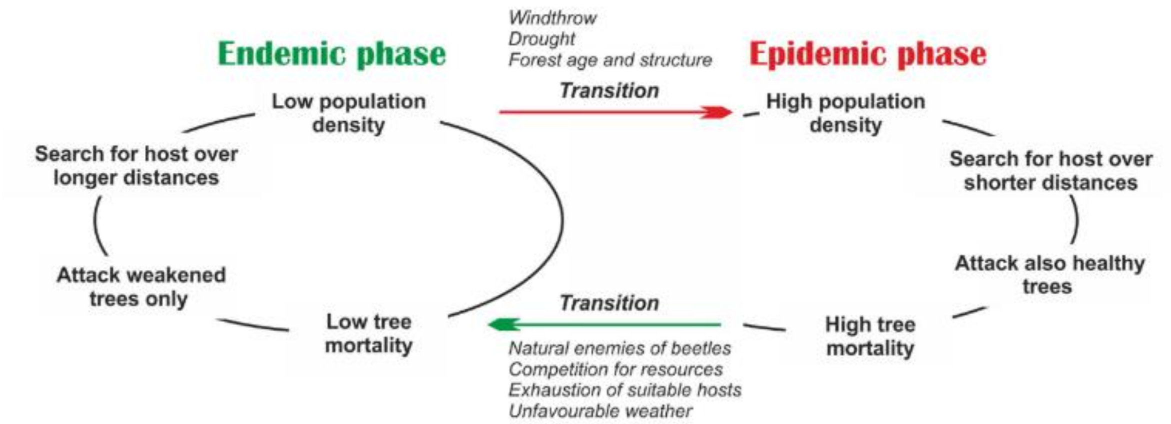


Figure 1. **Bark beetle population dynamic.** The diagram shows how the two phases of the population (endemic and epidemic) interact with each other. Factors that can trigger the phase change are indicated. Adapted from Hlásny et al. (2019).

1.2.1. Abiotic factors relevant for bark beetles

Abiotic factors relate to climate variables and create suitable conditions for the spread of bark beetles. In collaboration with other drivers, abiotic factors can cause great stress in the tree or even kill it (Millar & Stephenson, 2015).

For instance, extreme climate events (e.g., storms) can damage trees. The windthrows are a perfect habitat for the breeding of bark beetles since the tree is weakened and not able to generate opposition against the colonization of the insect (Hlásny et al., 2019). Another example relates to temperature. High temperatures can induce stress in the trees since a loss of water by evapotranspiration is generated, which decreases the capacity of the tree for defending against the colonization (Raffa et al., 2015). This, combined with drought conditions, can trigger the development of the bark beetle population (Millar & Stephenson, 2015). Noteworthy, the sun radiation received by the tree, and the temperature are determined by the topography of the area (Bentz & Jönsson, 2015).

According to previous research, the defence of the tree is influenced by the changes in water balance and carbohydrate content. The resin secretion is done by means of conduits, which are affected by the turgor pressure of the parenchyma cells. Particularly, the water content has an influence on these cells turgor pressure (Baier, 1996; Christiansen & Bakke, 1988).

One of the most important changes at a population-level occurs when weather conditions are very favorable for bark beetles. Then, the local population can turn from univoltine to multivoltine (Mezei, Jakuš, et al., 2017).

1.2.2. Biotic factors relevant for bark beetles

A planted forest is defined as "a forest predominantly composed of trees established through planting and/or deliberate seeding" (FAO, 2018). In Europe, great areas of Norway spruces have been seeded for obtaining timber resources due to its high productivity (Klimo, Hager, & Kulhavý, 2000). However, the biodiversity in this type of forest is usually much lower than in a natural forest, which increases the effect of natural disturbances (Hlásny et al., 2019). Therefore, planted forests are particularly prone to bark beetle infestation (Klimo et al., 2000).

As stated by Raffa et al. (2015), the forest structure has a strong importance in the outbreak. Specifically, a homogenous forest with mature individuals has a higher risk of being infested. On the contrary, heterogeneous forests are prone to maintain the population in the endemic phase (Hlásny et al., 2019). Not only the forest structure is a driver of the bark beetle outbreak, but also the tree diameter and bark thickness, which are directly linked to the number of offspring (Hlásny et al., 2019).

Moreover, as Jurc, Perko, Džeroski, Demšar, & Hrašovec (2006) stated, the location of the tree in the mountain can have an effect on the tree resistance against drought. The trees which are placed in the southern face of the mountain are adapted to more water-adverse conditions and develop a better root system, which makes them able to take water from higher depth in the soil. In addition, other biotic factors such as natural enemies, pathogens, symbionts, and competitors also have an effect on the population dynamic (Biedermann et al., 2019).

1.2.3. Bark beetles and forest management

The forest management approach followed in Europe tries to minimize the natural disturbances in order to maintain the forest structure and biodiversity (Seidl, Rammer, Jäger, & Lexer, 2008). Two main management strategies, according to the objectives of the forest owners and forest managers, can be differentiated (Hlásny et al., 2019). First, the multifunctional and production forest (MFPF) tries to reduce the bark beetle effects on timber production, since the main purpose is to get an economic benefit from the sale of wood (Hlásny et al., 2019; Zýval, Křenová, & Kindlmann, 2016). The ownership of this forest is normally private (Zýval et al., 2016). The second management approach occurs in high conservation value forests (HCVF). The management aims to maintain biodiversity and natural processes. Thus, the bark beetle outbreak is perceived as a natural disturbance with high importance on biodiversity, since it hits the monoculture forest, leaving space for regeneration with other tree species which will make the forest wilder (Hlásny et al., 2019; Lehnert, Bässler, Brandl, Burton, & Müller, 2013; Müller, Bußler, Goßner, Rettelbach, & Duelli, 2008). This kind of management normally takes place in state-owned forests (Hlásny et al., 2019; Zýval et al., 2016). The legislation differs in both types of forests. Compared to the MFPF, the HCVF has more legal restrictions related to the management of bark beetle (Hlásny et al., 2019).

Measures

The measures can be classified into consonance with the management strategy. For the MFPF, one of the measures is to reduce the rotation period, since the infestation is related to the age of the trees (Hlásny et al., 2019). Another strategy is to reduce the availability of host trees by supporting the increase of biodiversity in forests (Hlásny et al., 2019). The removal of trees that are affected by wind, snow, or ice also reduces the risk (Hlásny et al., 2019). For managing the bark beetle population, sanitation felling can be used. By the extraction of already colonized trees from the forest, an infestation of healthy trees can be avoided (Hlásny et al., 2019; Stadelmann, Bugmann, Meier, Wermelinger, & Bigler, 2013). Salvage logging is another technique that tries to diminish the effects of bark beetles on the timber quality and price by removing infested and damaged trees (Hlásny et al., 2019; Stadelmann, Bugmann, Meier, et al., 2013).

However, in the HCVF, different measures take place. The most important one is the zoning. With this procedure, a non-intervention area is left without intervention, fostering the natural development of the ecosystem and the biodiversity (Hlásny et al., 2019; Zýval et al., 2016). This area is surrounded by a buffer area in which the intervention is allowed on a different level, depending on the local management plan (Hlásny et al., 2019).

Monitoring bark beetle infestations

As aforementioned, most of the measures take place once the bark beetle is in the outbreak phase. However, forest management is aimed at minimizing the outbreak risk (de Groot, Diaci, & Ogris, 2019; Fahse & Heurich, 2011; Netherer & Nopp-Mayr, 2005; Overbeck & Schmidt, 2012). Knowing which areas are at higher risk is useful for the optimization of financial and labor resources (Netherer & Nopp-Mayr, 2005). This can be done by identifying local conditions that increase the outbreak probability (Pasztor, Matulla, Rammer, & Lexer, 2014).

Some efforts have been made to describe the population distribution of the bark beetle. In the review from Bentz & Jönsson (2015), it is concluded that the most important variables for describing the spread of bark beetles in every study change according to the area where the study was performed.

In this review, it could be demonstrated that in the Tatra mountains (Slovakia and Poland) various indicators, such as terrain, climate, soil, forest structure, or forest damage, are used in the description of bark beetle distribution (Netherer & Nopp-Mayr, 2005). Another study showed that in the European Alps, dry summers, and warm temperatures are the most important variables (Marini, Ayres, Battisti, & Faccoli, 2012). In Austria, it was demonstrated that climate means and extremes are the best indicators for the bark beetle outbreak, but they can be highly influenced by the forest management (Thom, Seidl, Steyrer, Krehan, & Formayer, 2013). In Switzerland, the most important variables for assessing the population dynamics were the temperature, volume of Norway spruce, and storm damage (Stadelmann, Bugmann, Wermelinger, Meier, & Bigler, 2013). Another study carried out in Sweden suggested that after a storm, damaged trees have a great influence on the bark beetle outbreak (Marini, Lindelöw, Jönsson, Wulff, & Schroeder, 2013). In the same area, the influence of different predictors studied, being the amount of Norway spruce the highest predictable effect (Kärvemo, Van Boeckel, Gilbert, Grégoire, & Schroeder, 2014). Finally, in Slovenia, it was revealed that the trees in the northeast slopes of the mountains are more exposed to bark beetle infestation since they are more sensitive to drought (Jurc et al., 2006). Another study in the same area suggested that the most important variables are the amount of Norway spruce trees, climate variables, salvage logging previous year, and the number of trees infested by bark beetles (de Groot & Ogris, 2019).

Peter Baier, Pennerstorfer, & Schopf (2007) created the PHENIPS model for simulating the brood development of the bark beetle. This model calculates the microclimatic conditions (temperature and solar radiation) by using the topography. It was validated in Kalkalpen National Park (Austria). The LPJ-GUESS ecosystem model was built based on the fact that in Sweden, the risk of infestation by bark beetle is directly related to the breeding material after extreme climatic events (Jönsson, Schroeder, Lagergren, Anderbrant, & Smith, 2012).

In the light of forest management, a model for simulating the required sanitary felling due to bark beetle infestation was created (Ogris & Jurc, 2010). This model used 21 different variables and was successfully applied in Slovenia. The impact of sanitation felling and salvage logging in Switzerland was analyzed and concluded that, if it is done at the right time, it can reduce the impact of the bark beetle in the forest (Stadelmann, Bugmann, Meier, et al.,2013). This idea is supported by further studies, which came to the conclusion that salvage logging has a more positive effect than sanitation felling (Havašová, Ferenčík, & Jakuš, 2017). Another agent-based model was developed by Fahse & Heurich (2011) for predicting the outbreak risk using the impact of antagonists and management. It was concluded that if 80% of beetles are removed, there is a low risk of having an outbreak.

1.3. Species distribution modeling

After the previously mentioned findings, it can be observed that every study uses a different statistical method for analyzing the importance of the variables in predicting a bark beetle outbreak. Some of the methods are principal component regressions (Thom et al., 2013), expert knowledge (Netherer & Nopp-Mayr, 2005), general linear models (de Groot et al., 2019; Marini et al., 2012), general additive models (Mezei, Blaženec, Grodzki, Škvarenina, & Jakuš, 2017),or boosted regression trees (Kärvemo et al., 2014).

All methods are species distribution models (SDM), statistical approaches that enable predicting the species occurrence based on environmental variables (Elith & Leathwick, 2009). Input data for these models are a dependent variable, which corresponds to the absence-presence data of a species, and the independent variables, which are the environmental predictors. The predictor variables can be numeric, binary, or categorical (Elith, Leathwick, & Hastie, 2008).

In former research, methodological obstacles have occurred. Conventional regression models have been shown to be too simplistic to represent interaction among variables (Elith et al., 2008). More methods were developed, but the problem came when every individual method was giving different results. Therefore, the concept of ensemble methods was introduced, with the purpose of using several models, providing more robust results (Araújo & New, 2007). This principle is used by boosted regression trees (BRT), a machine learning method that has been applied successfully in other investigations with the purpose of mapping species distribution (Akayezu, van Duren, Groen, Grueter, & Robbins, 2019; Froeschke

& Froeschke, 2011; Sproull, Bukowski, McNutt, Zwijacz-Kozica, & Szwagrzyk, 2017; Veran et al., 2016; Vorster et al., 2017).

Boosted regression trees

Boosted regression trees combine the predictions from weak classifiers - namely, simple regression tree models - with the purpose of getting a stronger classifier (De'ath, 2007; Elith et al., 2008). It uses two different algorithms for improving the performance: regression trees and boosting. Regression trees consists of continuous binary splits of the predictor variables with the purpose of fitting them to the response variable. Each of the groups in the end nodes is the most homogeneous possible (De'ath & Katharina, 2003). With this algorithm, the interaction between predictors can be modelled. The algorithm is very sensitive to the training data (Elith et al., 2008), but this can be solved by the use of boosting. This algorithm improves model performance by combining the results of different trees and getting the average (De'ath, 2007; Elith et al., 2008). This algorithm is sequential, which means that in every step, a new tree is built on the results of the previous tree, reweighting at every new stage and improving the performance (De'ath, 2007; Elith et al., 2008). Stochasticity is an important characteristic of this method, as it reduces overfitting and introduces variability in the results. This randomness is created in the model by only using a sub-sample of the training data for fitting every regression tree (De'ath, 2007; Elith et al., 2008).

1.4. Research problem and purpose of the study

Bark beetles cause damage in forests, where Norway spruces occur. Therefore, a map of areas that can be prone to bark beetle infestation is highly needed for decreasing the impact of this insect. Previous studies identified different variables that, in every region, determine the future presence of this species. In the Schwarzwald National Park, investigations regarding the mapping of areas prone to bark beetle infestation have been lacking. Therefore, the **objective** of the present study is to analyze which spatial variables can be linked to the outbreak of bark beetle in our study area (i.e., Schwarzwald National Park), in order to create an outbreak probability map. This research can be divided into three research questions:

1. What spatial variables are good candidates for building a predictive model for bark beetle infestation?
2. What statistical model provides the strongest prediction of bark beetle infestation?
3. Where is the probability of bark beetle infestation highest and lowest in the study area?

2. MATERIAL AND METHODS

The first part of this section describes the study area. Then, the structure follows the research questions (**Figure 2**). For the first research question, we want to identify the candidate variables for building the model. The list is explained in this section, as these variables will be needed for the selection of the variables in the second research question. Apart from that, the data and method used in the second research question are explained. This section concludes with an explanation of how the prediction map was done and the accuracy of the model calculated, which correspond to the third research question.

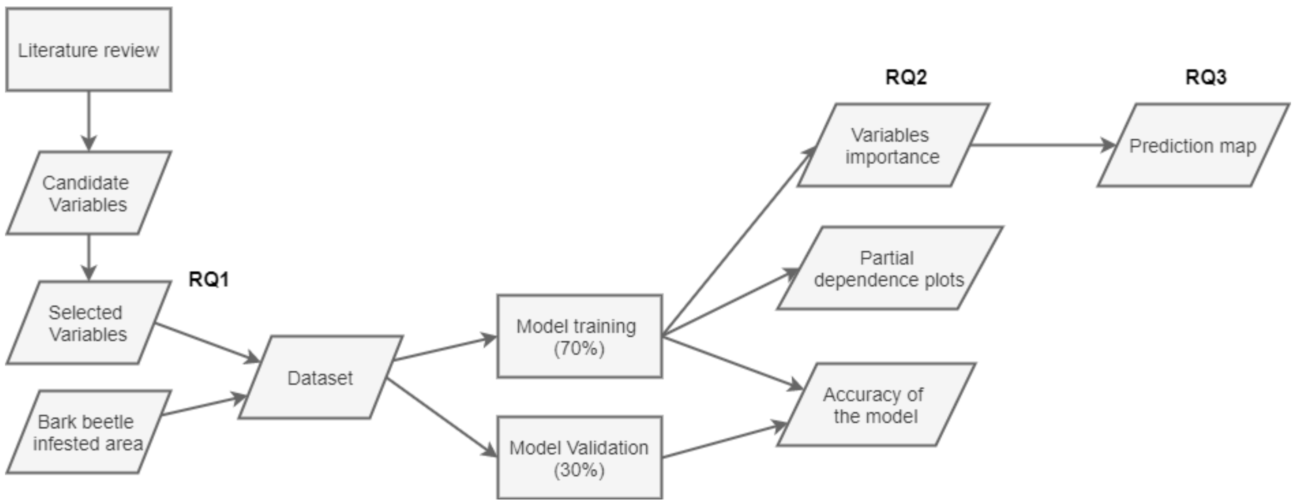


Figure 2. Steps of the method used in this research. This figure shows the steps taken in the methodology of the study and its relation with the research questions. The rectangles indicate a process and the parallelograms data inputs or outputs.

2.1. Study area

The research has been carried out in the Schwarzwald National Park, which is located in the south-west of Germany, in the federal state of Baden-Württemberg (**Figure 3**). In 2014, this area was declared National Park (*Waldmanagement im Nationalpark Schwarzwald*, 2017).

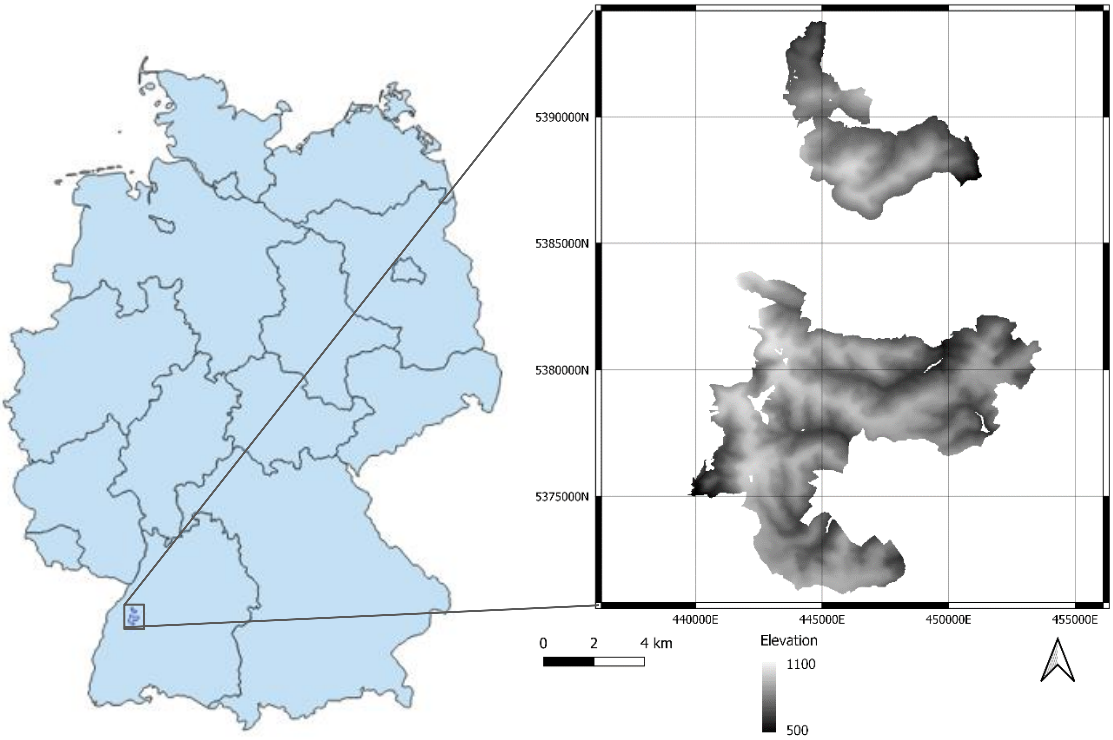


Figure 3. Study area and elevation map. The first image shows the location of our study area in Germany. On the right side, it is displayed a map of the elevation in the study area. This information was obtained from the Copernicus Land Monitoring Service.

The surface of the Schwarzwald National Park is 10,062 hectares. The elevation in the area ranges between approximately 500 and 1100 meters (**Figure 4**). Due to this variation in altitude, the temperature also oscillates, with a yearly average around 7.2 °C ("DWD portal. Freudenstadt weather station," 2020), with the highest temperatures reached within the months of June, July, and August. Conversely, in January and February, the lowest temperatures can be found (**Figure 4a**). The annual precipitation in the area is approximately 2100 mm ("DWD portal. Baiersbronn-Ruhestein weather station," 2020). The precipitation is well distributed over the year (**Figure 4b**). Part of this precipitation falls in the form of snow.

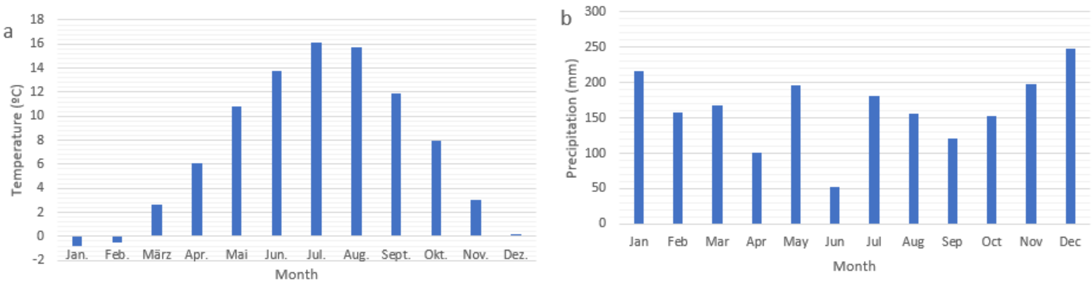


Figure 4. **Weather data of the Schwarzwald National Park.** The monthly average of the temperature (a) is from the period of 1981-2010, while the monthly average of precipitation (b) is from the period of 2006-2018. This information has been obtained from the Deutscher Wetterdienst portal.

The temperate forest in the National Park is dominated by Norway spruces (*Picea abies*), with 70% of the trees. The silver fir (*Abies alba*) is the second most dominant specie (12%), followed by Scots pine (*Pinus sylvestris*) (6%). The most frequent deciduous species is the European beech (*Fagus sylvatica*) (5%). (*Waldmanagement im Nationalpark Schwarzwald*, 2017). The bedrock is mainly granite and gneiss, originating gley and podzol soils, which are acidic and low in nutrients ("Black Forest National Park Portal," 2020).

The management in the National Park follows a zoning strategy (**Figure 5**) (*Waldmanagement im Nationalpark Schwarzwald*, 2017). 33% of the surface is part of the core area, where the natural processes cannot be altered by human influence. Forest management measures can only be taken if there is a critical disturbance that threatens the survival of a species, or for protection of neighbouring areas.

41% of the National Park area is designated as a development zone. This area will become part of the core zone after 30 years. The management in this area is done with the purpose of boosting natural processes. Examples of these measures are the increase of the mixed-forests area, restoration of bogs, and protection of species biotopes. In this area, bark beetles attacks in the endemic phase are considered important, since they create open areas for mixed-forest and drive the regeneration of the forest. However, there is a management plan against mass-attacks.

The remaining area (26%) is considered a management zone, in which human interventions have the aim of protecting surrounding areas. In this buffer area, intense bark beetle management is carried out, with measures such as weekly monitoring or sanitation logging. In order to increase the effectiveness of this area, in the long term, silvicultural measures aim to increase the mixed-forest cover.

The FVA institute is in charge of bark beetle monitoring. As soon as there is an outbreak alert, the committee "Regel-Jour fixe Borkenkäfer-Management im Nationalpark Schwarzwald" takes the decision about the management that is carried out. The stakeholders that are part of this committee are local authorities, representatives of the federal state government, FVA Institute, and the National Park managers (*Waldmanagement im Nationalpark Schwarzwald*, 2017).

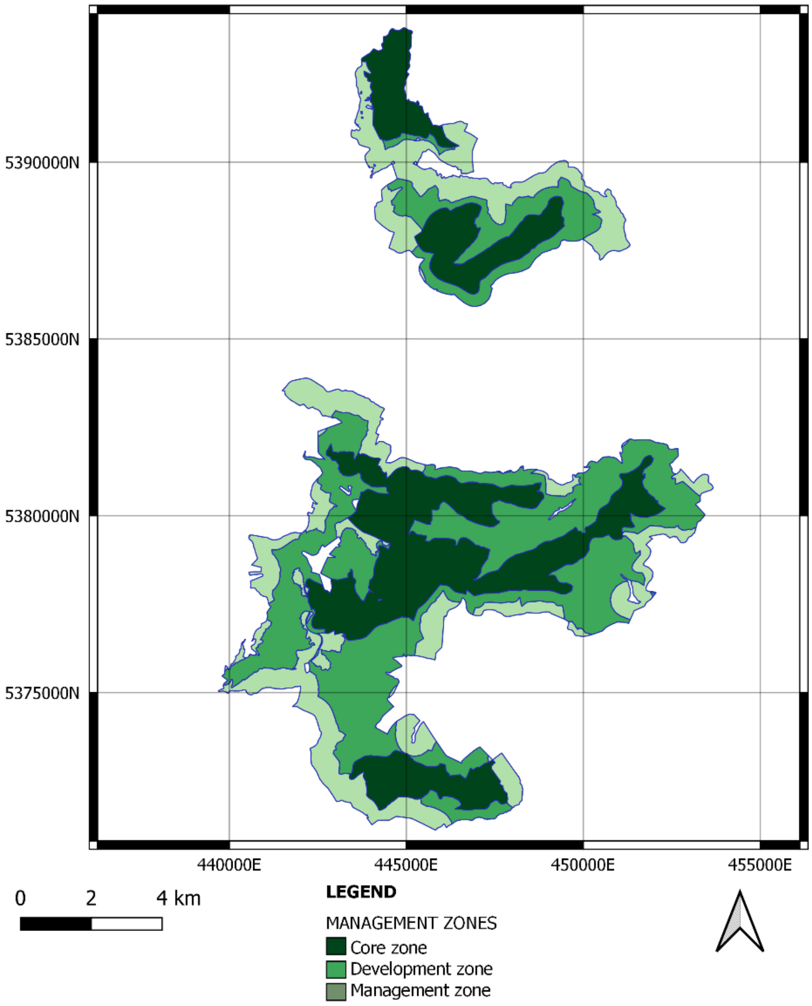


Figure 5. Management zones of the Schwarzwald National Park. In this figure, the different management zones are shown. This information has been provided by The Forest Research Institute Baden-Wuerttemberg (FVA).

2.2. Variables

A literature review was performed to answer the first research question "What spatial variables are good candidates for building a predictive model for bark beetle infestation?". Once all the variables that have been used by other authors were listed, the list was reduced, taking into account significance in the study area, and data availability. The variables can be grouped into four different groups: topography, vegetation, soil, and landscape.

2.2.1. Candidate variables

Topography

´The topographic variables were selected to examine how the location of the trees in the mountain could affect the infestation risk. As it is a mountainous area, the elevation and slope have high variability and potentially could be related to bark beetles presence (Sproull et al., 2017). Northness and eastness have been used in previous studies (Stadelmann, Bugmann, Wermelinger, & Bigler, 2014; Vorster et al., 2017) and are the transformation of the circular scale of the aspect to a linear scale. The irradiation (Baier et al., 2007) and topographic wetness index (TWI) (*FVA Annual report*, 2017) can explain the effect of sun and water in the distribution of bark beetles. TWI analyzes where the water is accumulated in the terrain (Mattivi, Franci, Lambertini, & Bitelli, 2019).

Vegetation

In relation to vegetation, three sub-groups of variables can be found. The first sub-group contains variables related to tree parameters, such as rooting depth, or leaf area index (LAI), which have been used by Netherer et al., (2019). Rooting depth influences the water storage capacity in the soil (Matthews et al., 2018). On the other hand, LAI is a variable that has an effect on transpiration and interception by the tree. The age of the Norway spruces was included by Seidl, Baier, Rammer, Schopf, & Lexer, (2007) since older trees are more prone to be infested. Norway spruces height can be taken as a representative of the age of the trees if the age of the tree is not available (Čihák & Vejpustková, 2018; Netherer et al., 2019; Thom et al., 2013).

The second sub-group of vegetation variables is related to forest structure. Among these variables, the number of Norway spruces and the total vegetation describe the tree canopy density per pixel (Netherer & Nopp-Mayr, 2005). The percentage of Norway spruces among all the tree species is indicative of the tree diversity of the forest (Thom et al., 2013).

Variables from the third sub-group are related to previous damages that have weakened trees, creating breeding substrate for bark beetles. One example of these variables is the number of Norway spruce, which were affected the previous year (de Groot & Ogris, 2019). A comparable variable is the sanitary felling data from the previous year (de Groot & Ogris, 2019; Kautz, Dworschak, Gruppe, & Schopf, 2011). Storm-felled trees of the previous year were also employed in recent studies (Marini et al., 2017; Netherer & Nopp-Mayr, 2005; Stadelmann, Bugmann, Meier, et al., 2013). In addition, Stadelmann et al., 2013 investigated salvage logging data, which is representative of damaged trees. Moreover, the snow effect on the trees was included in models with the parameter snow breakage (Netherer & Nopp-Mayr, 2005). The datasets from this sub-group can also be included in the models as distance maps.

Soil

The soil characteristics can determine the development of the roots (Puhe, 2003), which has an effect on the tree condition. Consequently, the soil also affects the defense system of the tree (Christiansen & Bakke, 1988). The depth of the soil is associated with the capacity to retain water (Hengl et al., 2017; Puhe, 2003), since deeper soils have a higher water retention capacity, which can contribute positively to the defenses against bark beetles (Rehschuh, Mette, Menzel, & Buras, 2017). Moreover, when the soil is shallow, sufficient vertical penetration of the roots is not possible. Thus, the tree has a higher risk of being affected by windthrow (Puhe, 2003). The soil organic carbon stock variable was previously used since it improves the soil structure and properties (Blanco-Canqui et al., 2013). Along the same lines, the cation exchange capacity (CEC) influences the tree access to minerals (Hobbie et al., 2007; Jentschke et al., 2001). Therefore, a high CEC could lead to a lower risk of bark beetle infestation, as shown by de Groot & Ogris, 2019. In the same article, soil base saturation was included as a representative of the calcium availability in the soil.

The type of soil, as well as some structural characteristics of the soil, such as porosity, bulk density, or permeability, were utilized by Ogris & Jurc, 2010. Netherer et al., 2019 further included the saturated water content of the soil derived from the soil texture.

Landscape

Regarding landscape variables, previous studies suggested that the higher the proximity to streams, the higher will be the water availability for the tree, hence the lower the infestation risk (Ogris & Jurc, 2010). By including the distance to paths and roads as a variable, it is possible to test whether the open areas can act as a corridor for the bark beetles spread. It is similar to the variable edge effect used by Netherer & Nopp-Mayr, 2005.

Apart from the groups that have been commented, previous studies implemented weather variables, such as temperature and precipitation (Baier et al., 2007; Marini et al., 2017). However, our area is not very extensive, which means that the spatial variation of how these variables affect the vegetation is rather related to the previously mentioned variables.

2.2.2. Selected variables

As the study area is a mountainous region, and the topography changes, all topographic variables have been selected. Regarding the vegetation variables, the selected predictors are Norway spruces height, number of Norway spruces per pixel, total vegetation per pixel, and the percentage of Norway spruces among all species in the pixel. The vegetation variables related to damage to the forest were not used because the model created does not take into account the temporal aspect. Rooting depth and LAI was not selected because the data was not available. The only soil variables available in the study area are soil depth, soil organic carbon stock, and cation exchange capacity. As landscape variables, both proximities to streams, and distance to paths and roads were used.

2.3. Data

Once the selection of variables has been made, the next steps required were data acquisition and data processing. These steps were different for every group of variables. The description can be found in this section.

2.3.1. Data acquisition

The boundary of the study area was obtained from the Department of Ecosystem Monitoring, Research, and Wildlife Conservation from the National Park. For the topographic variables, the digital elevation model (DEM) was downloaded from Copernicus Land Monitoring Service and had a resolution of 25 m.

The vegetation layer was also obtained from the Department of Ecosystem Monitoring, Research, and Wildlife Conservation from the National Park. This point layer was created with multispectral photography acquired in 2014/2015 and LIDAR data from 2015. A classification using random forest algorithm was carried out with trees higher than 15 m and a crown surface above 10 m^2. This shapefile contains information for every point about the monitoring date, tree species, height, and crown area.

The soil layers were downloaded from Soilgrids data portal. This portal has been created by ISRIC - World Soil Information. The layers have a resolution of 250 m. The datasets downloaded were depth to bedrock, in cm; soil organic carbon stock, in tons/ha and depths between 0-5 cm, and cation exchange capacity in the upper soil layer, in cmolc/kg. The landscape variables used in the model were obtained from the Department of Ecosystem Monitoring, Research, and Wildlife Conservation from the National Park.

Lastly, the bark beetle infestation layer was obtained from the National Park. This layer was created as part of the IpsPro-Project from the FVA Institute. In this project, aerial photography from 2014 to 2019 was used for classifying which areas suffer a bark beetle infestation in the Schwarzwald National Park. The used algorithm for classifying the images was random forest. Every polygon contains information about the infestation date.

2.3.2. Data processing

In this section, the steps taken for creating the input layers of the model are described. First, the variable layers, which were vectors, were all rasterized. At the end of this step, every layer had a pixel size of 25 m, and the coordinate reference system used was EPSG: 25832. The preprocessing of the datasets was performed using QGIS, and R (R Development Core Team, 2008). **Table 1** displays a brief description of the layers used. In the "Short name" column, the variable name for every layer during the processing steps is indicated. In the following, variable names in tables and figures will be referred to according to the short name provided in **Table 1**.

Table 1. **Description of input layers used.** The table provides information about the variables groups, the variables that belong to every group, the name used in the model, units of the maps, and range of values.

Variables group	Variable	Short name	Unit	Ranges
Topographic	Altitude	Altitude	m	490 – 1089
	Slope	Slope	degrees	0-90
	Northness	Northness	south or north	-1 to 1
	Eastness	Eastness	west or east	-1 to 1
	Irradiation	Irradiation	Wh/m2/day	2151-7106
	Topographic Wetness Index	Topographic Wetness Index		6.606–19.773

Vegetation	Norway spruces total	NS_sum	number of trees / pixel	1-22
	Norway spruces %	NS_perc	%	7.1-100
	Norway spruces height	NS_height	m	15-51.3
	Total vegetation	Veg_total	number of trees / pixel	1-36
Soil	Depth	Depth	cm	645-1523
	Soil organic carbon stock	Soil_org_stock	tons/ha	39-56
	Cation exchange capacity	Soil_cat	cmolc/kg	30-66
Landscape	Distance to paths	Dist_paths	m	0-285
	Distance to water sources	Dist_streams	m	0-863

Topography

The digital elevation model (DEM) was used for creating different layers (**Figure 6**). First, the altitude and slope were obtained. Second, the northness and eastness were calculated from the aspect, being northness the cosin, and eastness the sin of it. These maps have a value range from -1 to 1. The highest values indicate northern and eastern slopes, while lowest values indicate southern and western slopes, respectively.

The sun radiation was obtained with the r.sun.insoltime algorithm from GRASS (Hofierka & Súri, 2002; Neteler, Bowman, Landa, & Metz, 2012). The sun radiation was calculated for the 15th day of the month in each month in which the bark beetle is active (from April until September). Then, the average of all these maps was calculated. The unit of this map is Wh/m2/day.

In order to calculate the TWI, the Topographic Wetness Index from SAGA GIS (Conrad et al., 2015), implemented in QGIS, was used.

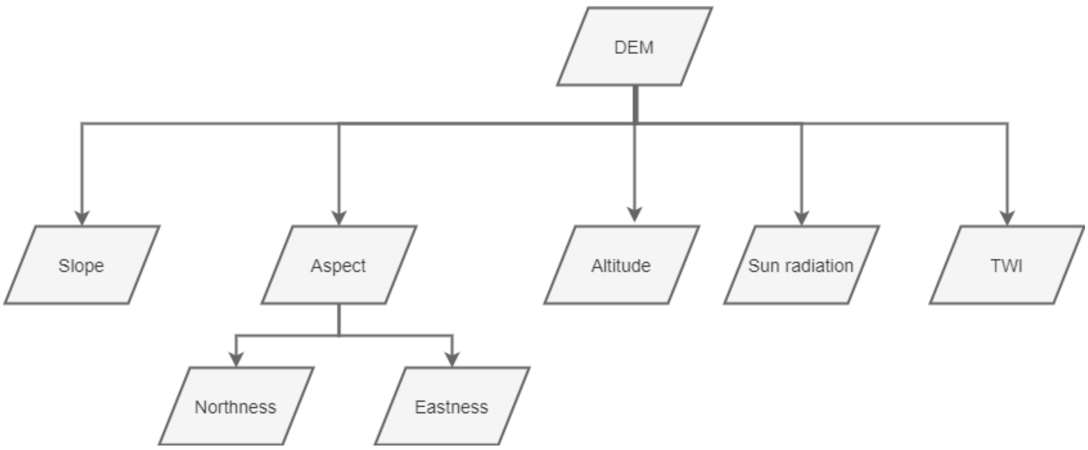


Figure 6. Workflow corresponding to the topographic variables. From the digital elevation model (DEM) different input maps were obtained.

Vegetation

Since the initial vegetation map was a point layer, the initial step was to rasterize this map in order to extract different information. First of all, the number of spruces per pixel was obtained. Then, the total number of trees per pixel was extracted. With the previous two maps, the percentage of Norway spruces compared to other species can be calculated. Finally, the average height of the Norway spruces per pixel was obtained. In **Figure 7** the processing steps taken for these layers are displayed.

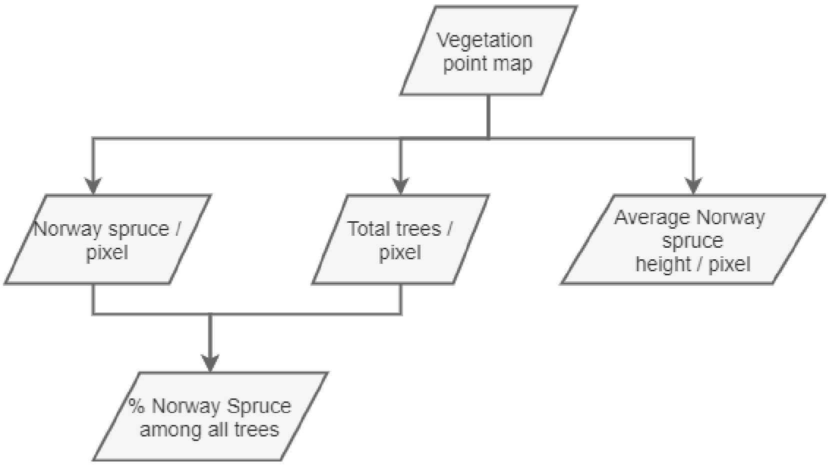


Figure 7. Workflow of corresponding to the vegetation variables. The vegetation map was rasterized, extracting different attributes.

Landscape

For the landscape variables, both maps followed the same processing steps. The layers were rasterized and reclassified. These maps were then used for calculating two raster maps with the distance to paths and roads, and distance to streams.

Soil

The only processing steps that the soil maps required were related to pixel size, coordinate reference system, and extent.

Infestation

For training the model, a layer with different sampling points indicating the presence or absence of the species is required. As the point layer of the Norway spruces was available, the sampling points selected correspond to an infested or non-infested tree. Regarding the presences sampling points, the first step was to overlay the infestation layer with the Norway spruces layer. We did not differentiate among years, because we were interested in the environmental conditions enabling infestations, rather than the temporal aspect. Randomly, 500 of the total infested trees were selected.

With respect to the absence sampling points, a buffer of 25 meters (one pixel) around the infestation area was created, to ensure that the environmental variables in presence and absence points were different. Subsequently, the area between the buffer and the National Park boundary was selected. This area was overlapped with the Norway spruces layer, and 500 of the non-infested trees were randomly selected. Finally, the presence and absence sampling points were merged in a layer that contained the 1000 observations. The processing steps can be seen in **Figure 8**.

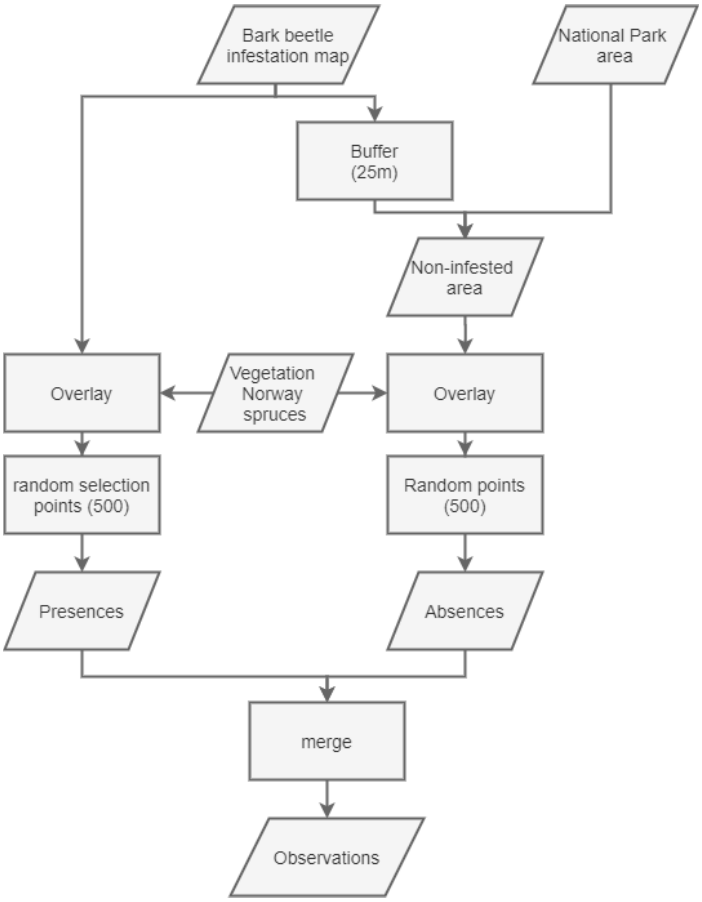


Figure 8. Workflow corresponding to the observations layer of the model. From the infestation area map, the vegetation map, and the National Park map, the observation layer with presences and absences has been obtained. The rectangles indicate a process, while the parallelograms indicate data inputs or outputs.

After preparing the dependent variable (presence or absence of bark beetles), and the independent variables (environmental variables), the environmental variables of every sampling point need to be extracted.

2.4. Boosted regression trees

Boosted regression trees are the method for answering the second research question "What statistical model provides the strongest prediction of bark beetle infestation?". Before creating the model, the collinearity needs to be analyzed, and related variables must be removed. Then, the model can be trained, and the weights and partial dependence plots can be created for results interpretation purposes.

Collinearity

Collinearity occurs when two or more variables follow a linear relation (Alin, 2010). Ecological models are sensitive to collinearity, as variables are normally dependent on each other to some extent. Collinearity can lead to prediction errors and, consequently, more difficulty in the interpretation of results (Dormann et al., 2013). In order to avoid the problem of multicollinearity among predictor variables, pairwise correlation coefficients, and variance inflation factor (VIF) were used. For these analyses, all observations (5000 presences and 5000 absences) were used.

For the pairwise correlation coefficients, Pearson's r-correlation indices were used, which analyzed the collinearity of every pair of variables (Dormann et al., 2013). With the purpose of visualizing the pairwise interactions, the "corrplot" package in R has been used (Wei, 2017).

The variance inflation factor (VIF) indicates how well one variable can be described by all the other variables. As a rule of thumb, VIF values higher than 10 indicated a collinearity problem among our predictors. Hence, when removing one of the variables with the highest VIF, the VIF of all the other variables should decrease (Naimi & Araújo, 2016).

After following these two approaches, the variables which had the highest collinearity were excluded from the model.

Model training and input parameters

As found in previous literature, 70% of the data was implemented for training the model (Akayezu et al., 2019). The rest was used for the accuracy assessment (**section 2.4.2**). The input data of the model consists of the dependent variable, which is the observations of bark beetles -or infested area-, and a set of independent variables, which are the predictor variables. The function applied for the model was "gbm.step" (Elith et al., 2008; Greenwell, Boehmke, & Cunningham, 2019), with Bernoulli response distribution, meaning that the distribution of the bark beetles was indicated using 0 for absences and 1 for presences.

There are different parameters that need to be selected for fitting the model. The first of them is the learning rate (lr). This parameter determines the contribution of every tree to the model (Elith et al., 2008). The tree complexity (tc) refers to the number of nodes in every tree. Lr and tc determine the number of trees (nt) required for the optimal prediction. As the lr is decreased, the nt will be increased, since each tree has a lower impact on the final model. For the tc, it can be observed that the more complex every individual tree, the less nt is required. The maximum nt that can be created is set to 10 000 (Elith et al., 2008). The last parameter to select is bagging fraction (bf), which adds stochasticity to the model by taking random subsamples of training data for each iteration (De'ath, 2007). This parameter indicates the percentage of data that will be selected at each iteration.

The optimal values for the parameters are those whose combination results in the lowest predictive deviance (Elith et al., 2008). After some combinations of the parameters, the best combination for fitting the model was lr=0.01, tc=5, and bf=0.5. The model, using a 10-fold cross-validation method, generated 1500 trees.

The 10-fold cross-validation is a method used for estimating the expected, predicted error (PE) (Hastie, Tibshirani, & Friedman, 2001). It consists of splitting the data into ten subsamples equally sized and using 9 subsets for training the data. One of the subsets is used for the validation (De'ath, 2007; Naimi & Araújo, 2016). The model has been fitted using the "gbm" package (Greenwell et al., 2019) in R (R Development Core Team, 2008).

Aggregated Boosted Trees

Between some of the presences, high proximity was detected, which was creating a sampling bias problem (Komori, Eguchi, Saigusa, Kusumoto, & Kubota, 2020). In order to correct this, as well as to avoid model overfitting, smaller sub-samples of the dataset were selected, creating a collection of 10 BRT models. After running them, the predictions have been aggregated by calculating the average. This variation of BRT is called aggregated boosted trees (ABT) and was proposed by De'ath (2007). Thus, the only difference in every model was the observation layer. The step of the random selection of points (**Section 2.3.2**) for creating the observation layer was repeated for every sub-sample. Each of the sub-samples was composed of 500 presences and 500 absences.

Weights

The contribution of each variable in the response is calculated with the number of times that the variable was used for splitting in the model, multiplied by the square of the weighted improvement of the model in every split (Friedman & Meulman, 2003).

Partial dependence plots

Partial dependence plots help to visualize the dependence between response and individual predictors (Friedman, 2001). The data were extracted with the "plot.gbm" function, from the "gbm" package (Greenwell et al., 2019). Subsequently, the mean was calculated for all models, and the "ggplot2" package was used for displaying the results (Wickham, Chang, Henry, Lin Pedersen, & Takahashi, 2020).

2.5. Prediction map and accuracy assessment

To answer the third research question, "Where is the probability of bark beetle infestation highest and lowest in the study area?" a prediction map was made using the prediction model accompanied by an accuracy assessment. The prediction map has been created with all the variables used for model training. To that aim, the fitted functions created by every sub-model were implemented in every pixel of the study area, obtaining ten prediction maps. Then, a new map with the average of all the sub-models prediction maps was created.

For the validation of the model, different methods were selected. Since the output of the model is a probability of species occurrence, the first step for carrying out the assessment of the model is to select a threshold, from which all pixels above will be classified as presence. The maximum value of Kappa in every model was selected as a threshold (Liu, White, & Newell, 2013). Once the predictive presence or absence of the species was obtained, the error matrix was created (Fielding & Bell, 1997). In the confusion matrix, predicted and observed values are compared (**Table 2**), extracting the number of true positives (a), false positives (b), false negatives (c), and true negatives (d) (Allouche, Tsoar, & Kadmon, 2006). From these values, different predictive accuracy measures can be calculated. 30% of the data was used for the purpose of accuracy assessment.

Table 2. Error matrix. It was adapted from (Fielding & Bell, 1997).

		Validation dataset	
		Presence	**Absence**
Prediction	**Presence**	True Positives (a)	False Positives (b)
	Absence	False negative (c)	True Negative (d)

AUC

The receiver operating characteristic curve (ROC) assesses how good the model discriminates between presences and absences. In this graph, the false positive rate (1-specificity) is plotted in the x-axis against the true positive rate (sensitivity) in the y-axis. A quantitative index for assessing the ROC is the area under the receiver operating characteristic curve (AUC). It ranges from 0.5 to 1, being considered 0.5 as a random model without accuracy, and 1 as a perfect discrimination capacity model. The higher the value, the closer the curve to the left top corner of the ROC graph (Hanley & McNeil, 1982).

TSS

True skill statistic (TSS) contrasts the correct predictions, minus those that can be random guessing, against a hypothetical set of perfect forecasts (Allouche et al., 2006). In order to calculate TSS, first, the sensitivity and specificity need to be calculated (**Table 3**). The sensitivity corresponds to the proportion of presences that have been correctly predicted. In other words, it expresses the omission errors. On the contrary, specificity is the probability that an absence will be correctly predicted, indicating the commission errors (Allouche et al., 2006). The scale of the TSS ranges from -1 to 1, being values under 0 considered a random model and values close to 1, indicating perfect agreement (Allouche et al., 2006).

Table 3. Formulas for sensitivity, specificity, and TSS. Adapted from (Allouche et al., 2006)

MEASURE	FORMULA
SENSITIVITY	$\frac{a}{a+c}$
SPECIFICITY	$\frac{d}{b+d}$
TSS	Sensitivity + specificity - 1 = $\frac{ad-bc}{(a+c)(b+d)}$

For analysing the spatial variability of results among the different sub-models, an uncertainty map has been created. This has been done by calculating the standard deviation of the sub-models results per pixel.

3. RESULTS

In this chapter, the data correlation is shown. Then, the results of the model, including weights, the data distribution of the variables, and partial dependence plots are presented. The last sections include the display of the prediction map and the results of the accuracy assessment.

3.1. Correlation analysis

Figure 9 displays the correlation between the different variables. The variables that have the highest correlation (0.801) are the number of Norway spruces and the total vegetation per pixel, being the latter excluded from the input data of the model due to collinearity, (see discussion). The amount of Norway spruces was also positively correlated (0.559) to the percentage of Norway spruces in comparison to the other vegetation. Besides that, we observed a negative correlation between eastness and irradiation (-0.732). Apart from this pair of variables, we detected a negative correlation (-0.505) between slope and irradiation, as well as between altitude and NS height (-0.553). It is noteworthy that, apart from the previous correlation mentioned, the altitude variable was correlated, on a lower intensity, with 11 of the other variables.

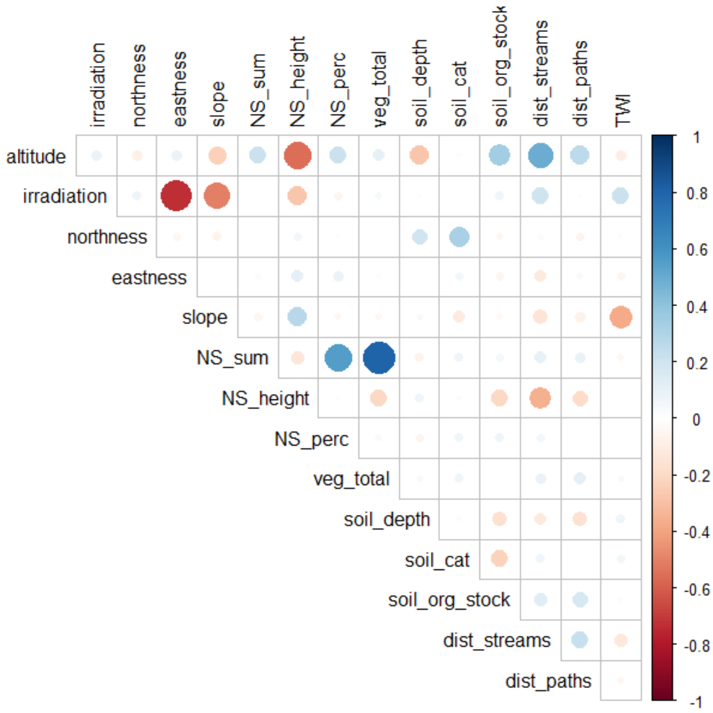


Figure 9. Correlation analysis among the variables. The bigger the size of the circle, and the intensity of the color, the bigger the correlation between the two variables. Red circles indicate a positive correlation and blue circles a negative correlation. Only values below -0.1 and above 0.1 are displayed. This correlation matrix has been done using 10.000 observations.

In **Table 4**, the values of the variance inflation factor (VIF) are displayed. In the second column of the table (VIF 1), the VIF values of all the variables are listed. The amount of Norway spruces, and the total vegetation show the highest VIF. It was decided to remove the latter from the input variables since we consider that the first one has a more predictive performance. After removing the variable "total vegetation", it can be observed that the rest of the VIF values are also lowered (VIF 2 column in **Table 4**).

Table 4. VIF values for every variable.

Variable	VIF 1	VIF 2
Altitude	2.301	2.275
Irradiation	5.045	5.045
Northness	1.200	1.198
Eastness	3.735	3.733
Slope	2.585	2.585
NS_sum	13.016	1.517
NS_height	1.7144	1.662
NS_perc	4.614	1.546
Veg_total	9.091	-
Soil_depth	1.176	1.176
Soil_cat	1.233	1.233
Soil_org_stock	1.255	1.253
Dist_streams	1.449	1.447
Dist_paths	1.126	1.125
TWI	1.248	1.248

3.2. Model results

From the model, different outputs were obtained. First of all, the relative importance of every variable in the model. Besides, the partial dependence plots of the six most influential variables were created. For a better interpretation of the results, the data distribution of the most influential variables is also shown.

3.2.1. Weights

The boosted regression tree analysis resulted in the relative importance of different variables in predicting the occurrence of bark beetle. This is expressed in **Table 5** as a weight percentage. The number of Norway spruces with 19.19% of the weights (**Table 5**), had the most significant influence. This variable is followed by the height of the Norway spruces and the altitude, with weights between 10-11%. The variables soil depth, slope, and percentage of Norway spruces have weights between 6% and 8%.

Table 5. Weights of every model variable expressed in %.

Variable	Weight
NS_sum	19.19
NS_height	10.87
Altitude	10.11
Soil_depth	8.51
Slope	7.93
NS_perc	6.92
Dist_streams	5.84
TWI	5.58
Eastness	5.45
Irradiation	5.26
Northness	4.80
Soil_cat	3.80

Soil_org_stock	2.94
DIst_paths	2.72

3.2.2. Data distribution of the most important variables

In **Figure 10,** the data distribution of the most important variables in our model are shown and was visually inspected for normal distribution. For the NS_sum variable, most of the observations are pixels with only one Norway spruce. The pixels with less than 11 Norway spruces have a similar high frequency. After that, the frequency gradually decreases, being the maximum 28 Norway spruces per pixel.

The height of the Norway spruces follows a normal distribution, being the highest frequency on 25 m. The range of this variable is between 15 and 49 m. The altitude variable is also normally distributed. Elevation values from 511 to 1133 are visible, but most of the observations are found between 700 and 1000 m. In the soil depth variable, we can also observe a normal distribution of the values. Most of the data are between 700 and 1100 cm, being 644 and 1528, the lowest and highest values. With regard to the slope, a normal distribution of the values is shown, while most of the observations have a lower slope than 30 degrees. For the percentage of Norway spruces, we observe that the number is increasing, until reaching the maximum, which is 100% of Norway spruces.

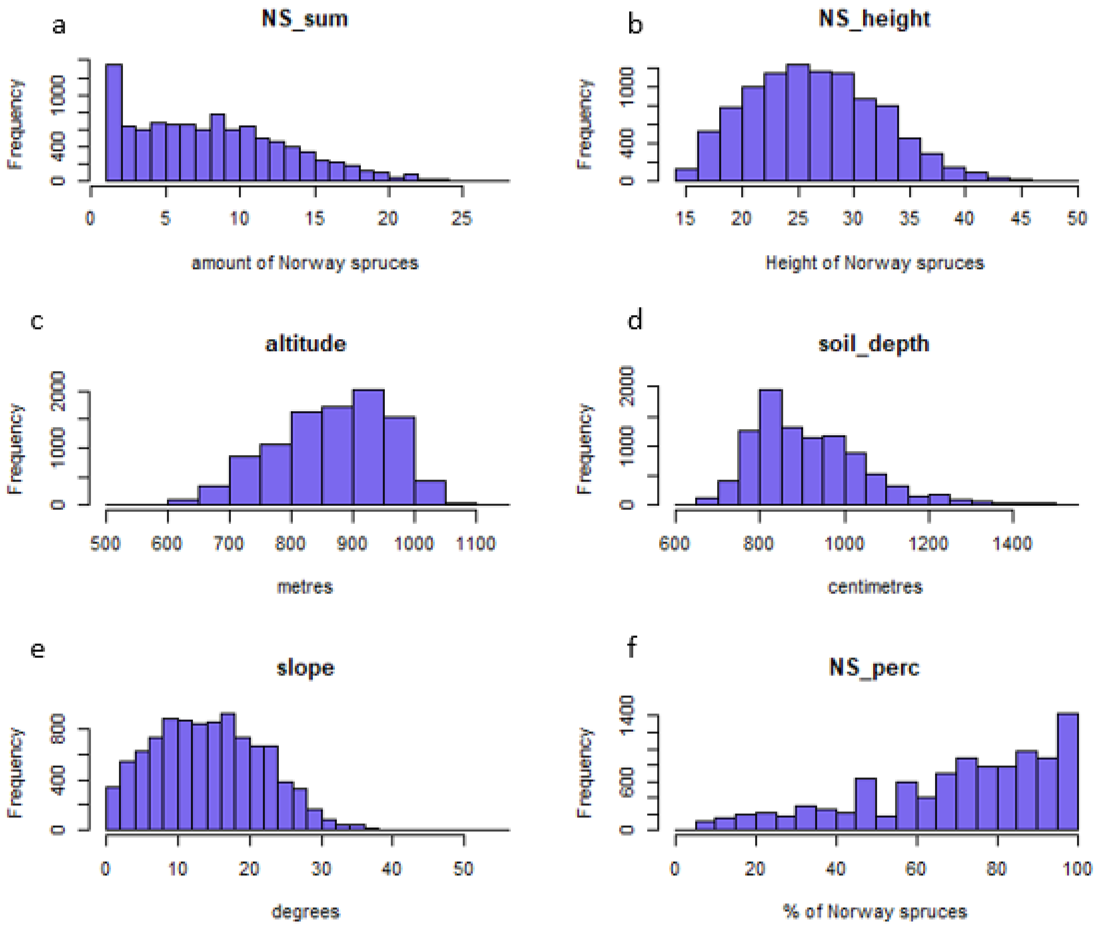


Figure 10. Data distribution of the most influential variables. The variables are the number of Norway spruces, the height of Norway spruces, altitude, soil depth, slope, and percentage of Norway spruces among all the tree species.

3.2.3. Partial dependence plots of the most important variables

The partial dependence plots facilitates the interpretation of the effect of every variable in the model (Elith et al., 2008; Friedman, 2001). In **Figure 11**, the partial dependence plots of the six most important variables in the model are displayed.

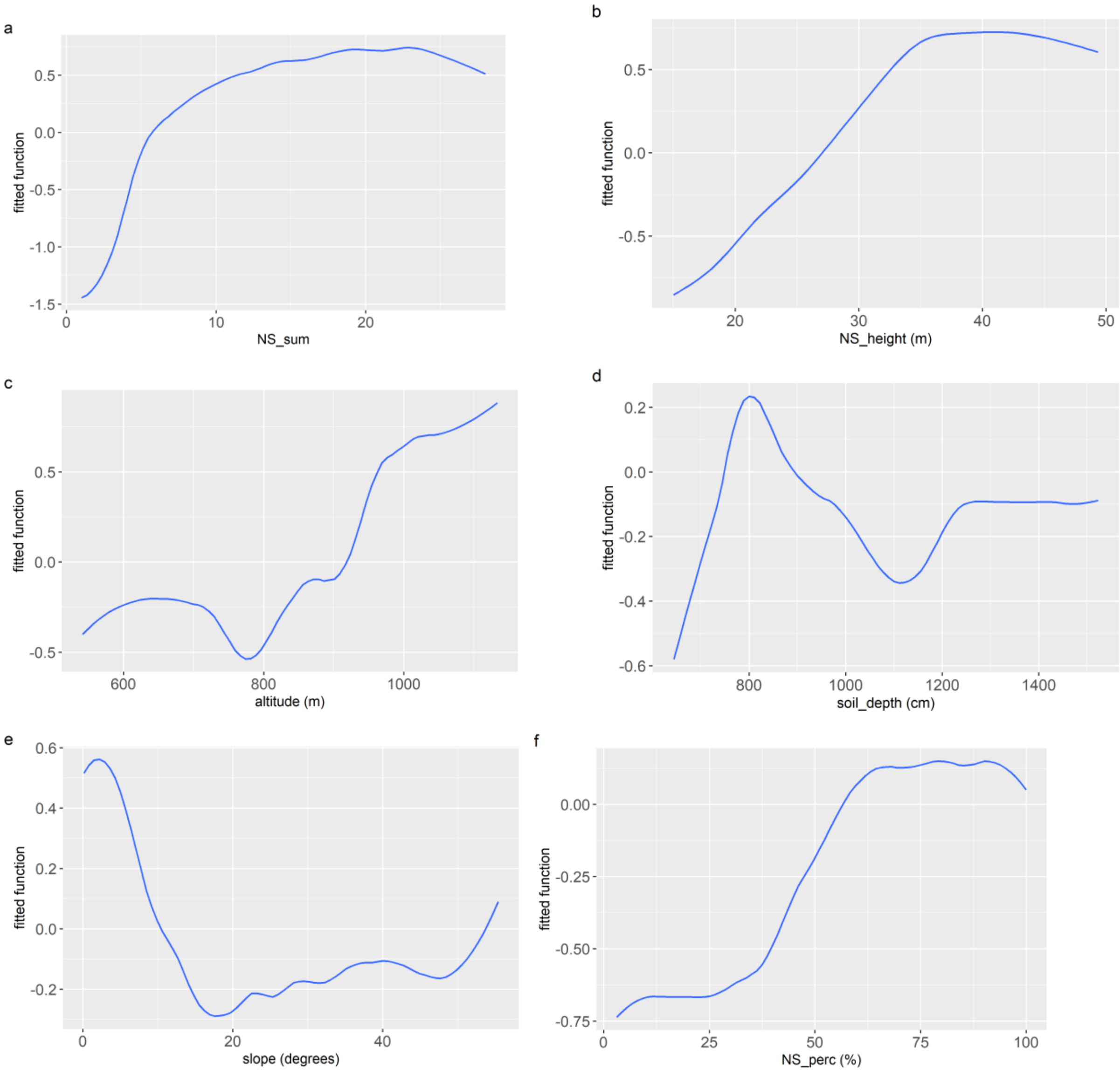


Figure 11. Partial dependence plots of the variables with more importance in the prediction. The y-axis shows the change in the fitted function, and the x-axis shows the range of values for every variable.

Figure 11a illustrates that the more Norway spruces are found in a pixel, the higher infestation probability the trees in this pixel have. In contrast, pixels with values under 5 have a very low probability of being infested. Furthermore, with a presence of more than 21 Norway spruces per pixel, there is a drop in the curve. The shape of the curve for the Norway spruce height variable (**Figure 11b**) is gradually increasing, being the older trees the ones which have the highest probability of being infested. For the altitude (**Figure 11c**), we see that the higher areas have more probability of bark beetle infestation. Between 700-850 m, a drop in the curve indicates that the probability of infestation decreases. **Figure 11d** illustrates that trees located on shallower soils are more prone to be infested. The curve is growing at the beginning until it reaches a peak at 800 cm depth. Then, it drops until its lowest value to 1100 cm, until it is growing again.

We must mention that the change in this variable response is not that big, as it ranges from -0.4 to 0.2. As displayed in **Figure 11e**, flat areas are more prone to be infested. The curve suffers a drop from the lowest values. After 10 degrees of slope, the probability starts being negative. Between 20-40 degrees, it stays constant, before dropping with slopes higher than 40 degrees. Regarding the percentage of Norway spruces variable (**Figure 11f**), the risk is very low at the beginning of the curve, but increases after 25% and around 60% become positive, indicating that spruce-dominant forest is more prone to be infested.

3.3. Prediction map

Figure 12 displays the bark beetle occurrence probability in the Schwarzwald National Park with the variables used in this study.

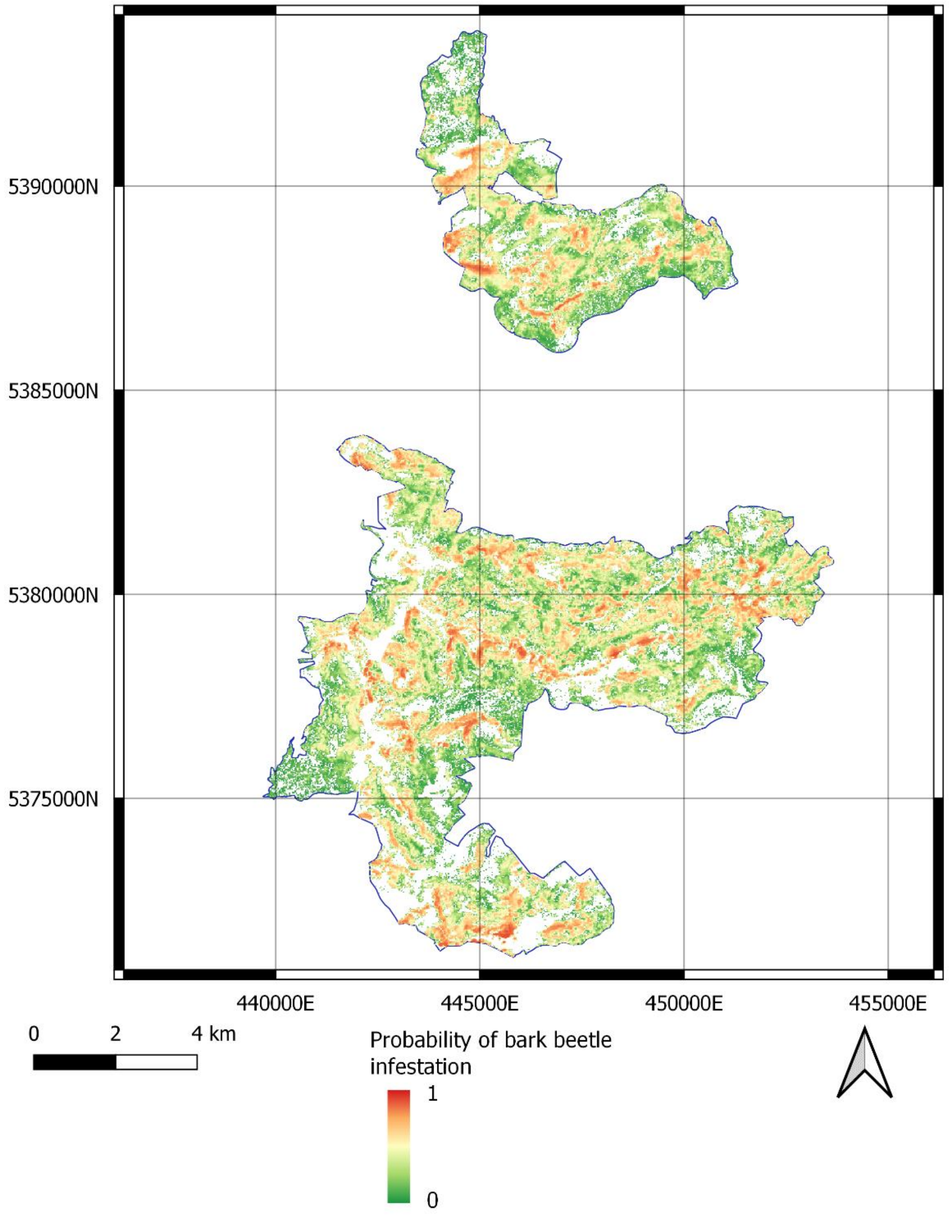


Figure 12. Bark beetle infestation probability map. It has been done getting the average value from the ten sub-models. For creating the sub-models predictions, all variables have been used.

3.4. Accuracy assessment

In **Table 6**, the accuracy of each of the ten sub-models is displayed. The average area under the receiver operating characteristic curve (AUC) of all models is 0.80. The third and sixth sub-models, with an AUC of 0.782 and 0.783, respectively, were the least accurate. The sub-model with the highest AUC was number 9, having an AUC of 0.856.

Table 6. Accuracy measures for the different sub-models. AUC refers to the area under the receiver operating characteristic curve and TSS to true skill statistic.

Sub-model	AUC	Sensitivity	Specificity	TSS
1	0.831	0.847	0.727	0.574
2	0.793	0.834	0.647	0.481
3	0.782	0.778	0.664	0.442
4	0.805	0.743	0.708	0.452
5	0.838	0.775	0.783	0.558
6	0.783	0.836	0.628	0.464
7	0.815	0.743	0.755	0.498
8	0.808	0.739	0.780	0.519
9	0.846	0.932	0.615	0.546
10	0.793	0.760	0.702	0.462

When using the TSS as a validation method (**Table 6**), the sub-models register an average TSS of 0.49. The sub-model 1 with a TSS of 0.574 has the highest accuracy, while sub-model 3, with 0.442, has the lowest TSS. The average value for sensitivity was 0.79, and for specificity 0.70.

For visual inspection of the model performance, the prediction map has been overlaid with the infestation map (**Figure 13**). Most of the large infestation locations are in the areas with higher outbreak probability. However, it can be observed that some of the infested areas are also in areas with low outbreak probability.

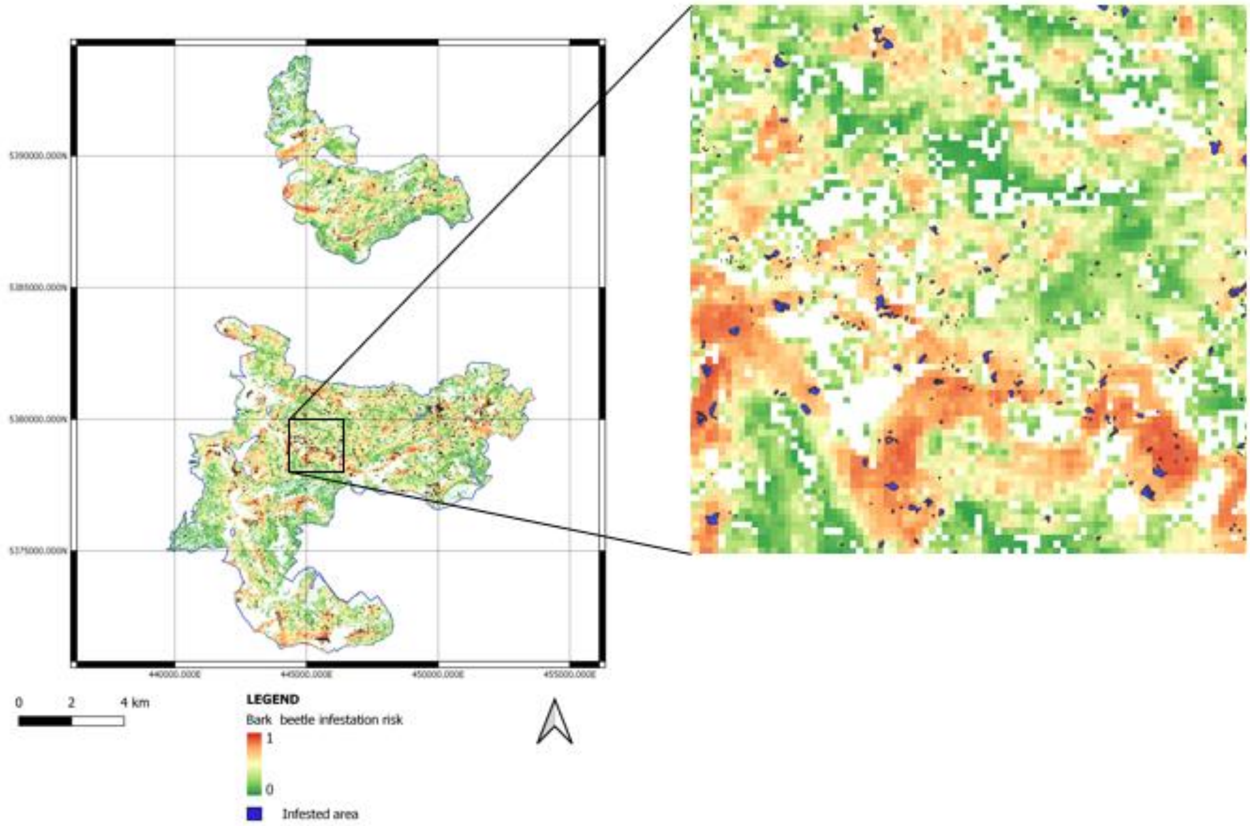


Figure 13. Outbreak probability map and infested areas in the Schwarzwald National Park. The prediction is the output of the model built with spatial variables, and the infestation layer has been created with an aerial survey.

Figure 14 shows how uncertainty is distributed over the area. The range of uncertainty values is from 0 to 0.22. Lower values indicate that the predictions of the sub-models were similar.

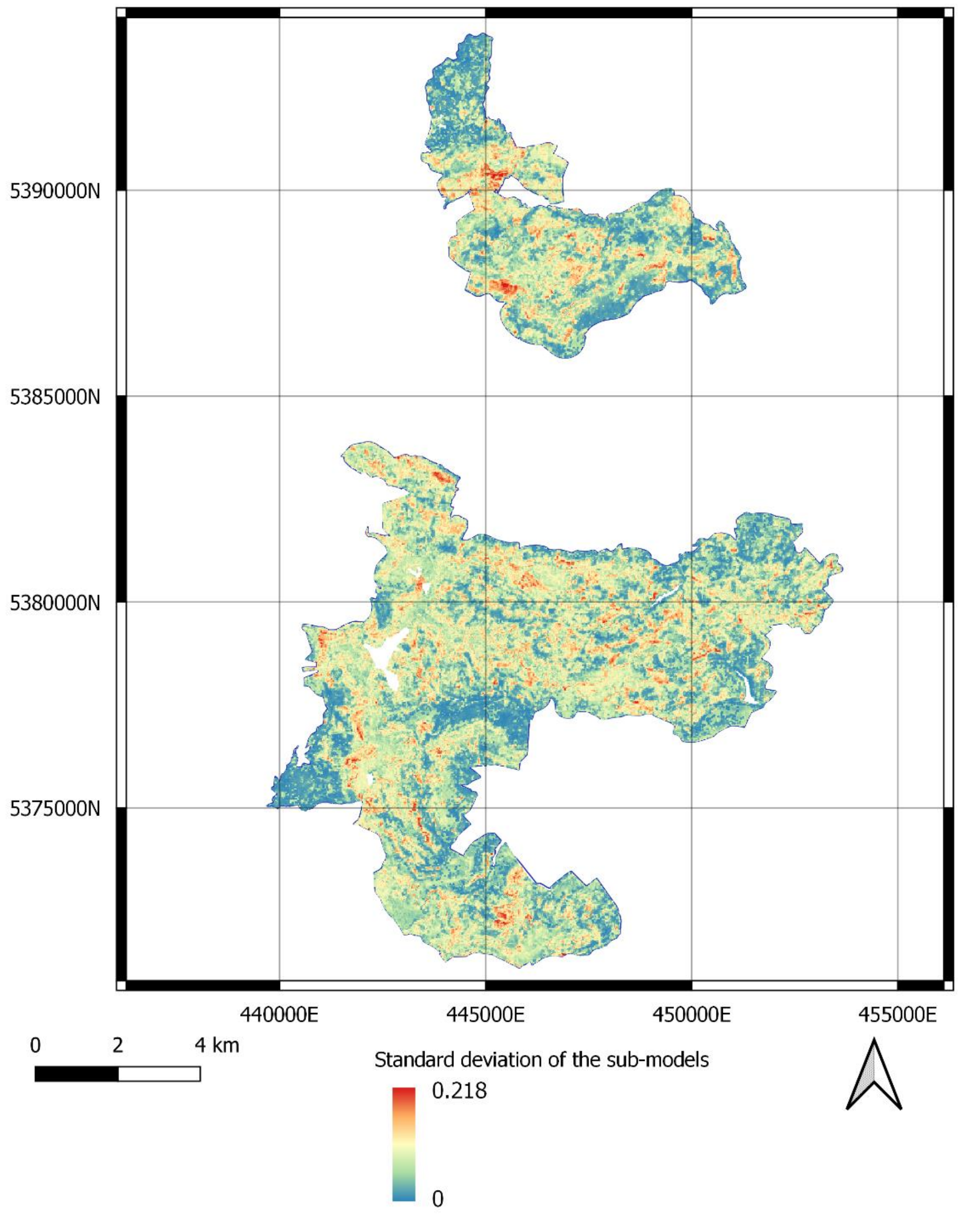


Figure 14. Uncertainty map. It shows the standard deviation of the sub-models results per pixel.

4. DISCUSSION

In the present thesis a model was calculated which helps to foster the understanding of variables that can be used for describing bark beetle outbreaks. In the following chapter, the performance, weaknesses, and strengths of the model are presented. In the last section, the findings will be discussed in the light of previous investigations.

4.1. Performance of the bark beetle infestation model

Overall, using the area under the receiver operating characteristic curve (AUC) for assessing the performance, all sub-models had, at least, a fair accuracy, as they are above 0.78. According to Araujo, Pearson, Thuiller, & Erhard (2005), an accuracy of 0.8 is considered as a good accuracy. Moreover, the model has also been validated with true skill statistic (TSS). Using the TSS method, the results were positive (0.48), indicating accuracy of the model (Allouche et al., 2006). While AUC is a threshold-independent measure (i.e., it tests all possible thresholds), TSS is dependent on the selected threshold (Allouche et al., 2006). As recommended by Allouche et al., 2006, both methods have been used, namely, a threshold dependent and a threshold independent measure. When comparing sensitivity and specificity, the latter one was higher, indicating that in the present model, the commission error was bigger than the omission error. The reason why kappa was not used in the current study is its proven dependence on prevalence (Allouche et al., 2006).

The uncertainty map (**Figure 14**) illustrates the spatial reliability of the results, taking into account the different results of every sub-model. In combination with the outbreak probability map, the uncertainty map should help to prioritize the monitoring areas for forest managers. In **Figure 13**, a visual interpretation of the results and the infested areas can be seen. The infested regions in areas with low outbreak probability might be explained by the micro-conditions that every tree has. Improving the data quality could enhance the accuracy of the model, as discussed in the next sub-section.

Data

In order to enhance model performance, some considerations for the input data can be taken. The vegetation layer used for creating the vegetation input layers represents only the trees, which are higher than 15 m. Bark beetles need a minimum of bark thickness of 2.5 mm to create the burrows (Grunwald, 1986). Since the bark thickness and height are related, modeling this relation can help to select a threshold. Importantly, this model should be applied locally, as it varies between regions (Stängle, Sauter, & Dormann, 2017).

Furthermore, for every sub-model, 1000 observations were selected. Future studies could further explore this issue by testing how the sample size can have an effect on the results of the model. This has been done previously by Elith, Leathwick, & Hastie (2008).

Due to data availability, the absence observations were pseudo-absences. This means that it is assumed that wherever there is not a presence, there is an absence (Naimi & Araújo, 2016). Nevertheless, this assumption needs validation, as some pseudo-absences might not be true absences. These observations were obtained outside a buffer of 25 m from the infested area. Further studies could focus on assessing the sensitivity of the buffer size in the model.

With regard to the soil layers, its coarse resolution (250 m) is likely not suitable for this study area. In mountainous areas, the soil development processes are influenced by the topography, with a very high spatial variability of soils at a local scale (Egli & Poulenard, 2016; Kruk, Ryczek, Klatka, & Malec, 2018).

Therefore, soil layers with better resolution should be used. Likewise, the sensitivity of the model to a 25 m spatial resolution should be tested.

Model

Moreover, some changes in the model training and model accuracy could have an impact on model performance. As mentioned previously, TSS is a threshold-dependent statistic. This type of method is not completely accepted, as it is dependent on the previously selected threshold, which can cause an overestimation of the true positives (Stephanie, Ceri, & S.J., 2001). For the present model, the maximum value of kappa was selected as a threshold. However, other thresholds should be tested for accuracy assessment (Freeman & Moisen, 2008; Liu et al., 2013).

For this study, the dataset has been partitioned, using 70% of the data for model training and 30% for validation, as it was done by Akayezu, van Duren, Groen, Grueter, & Robbins (2019). Nevertheless, it is always difficult to find the right balance, as increasing one threshold will negatively affect the other one. This issue should be further examined by future research.

The first idea of the current investigation was to create only one model. However, by having more data samples in the same area, the distance between each other is reduced, provoking spatial autocorrelation (Veloz, 2009). With the aim of avoiding spatial autocorrelation, different sub-models were created. Subsequently, the mean of all of them was calculated and used for further analysis. In the future it should be tested, whether the modification of the number of submodels leads to different results.

Correlation between variables

In the same way, the correlation between variables can affect model performance. As has been shown, the total amount of vegetation and the amount of Norway spruces are highly correlated. This can be explained by the great percentage of the vegetation (70%) composed by Norway spruces. The positive correlation between the number of Norway spruces and the percentage of Norway spruces per pixel demonstrates that Norway spruces are mainly located in low diversity areas. The other positive correlation that has been found was between distance to streams and altitude. This may relate to the location of the streams in the lowest areas of the National Park. Therefore, in pixels with low altitude, the distance to the streams is also lower.

Regarding the negative correlation of the variables, three main relations can be observed. First of all, the irradiation and eastness are highly correlated, implicating that the west faces of the mountains receive more irradiation. Another correlation that can be observed is between altitude and the height of the Norway spruces. A possible explanation might relate to the availability of resources. Norway spruces originate from North Europe, regions where the precipitation is higher and the sun radiation through the year lower. In the Schwarzwald National Park, higher elevation is related to higher sun radiation and lower water availability, possibly mediating as growth limiting factors. Consequently, as the tree mortality in elevated areas is higher, younger trees have the opportunity to grow, thus, boosting the regeneration of the forest. In addition, a negative correlation was revealed between irradiation and slope. Notably, the top of the mountains of the National Park is not very steep (see **Figure 3**). As a consequence, higher areas, which also are the ones that receive more sun radiation, have a lower slope.

4.2. Interpretation of the most important prediction variables

The partial dependence plots aid in the interpretation of the variable responses. These results, together with previous studies, can provide a basis for understanding the ecology of bark beetle outbreaks. When interpreting the partial dependence plots, it is important to consider the data distribution. A communality of most variables partial dependence plots is that the curve is more sensible in values that have a smaller number of observations, resulting in unexpected probability.

Number of Norway spruces

As **Figure 11a** shows, the more Norway spruces are in an area, the higher is the probability of finding bark beetles. This result ties well with the study from Kärvemo, Van Boeckel, Gilbert, Grégoire, & Schroeder (2014). The explanation seems to be related to tree species diversity. As mentioned before, areas with higher Norway Spruce density are mainly areas with low species diversity. According to Bauhus et al. (2017), the mixed forests resist better against specialized insect pests, as the density of feeding resources are reduced, and the insect population cannot evolve to the epidemic attack. In the same article, it is argued that in mixed forests, the host trees are more difficult to reach due to the distance that separates them. Another explanation for the lower infestation probability in mixed forests has to do with how the bark beetle colonizes trees. In this process, the first bark beetle arrives at a susceptible tree, and, if the tree is weak, it segregates a pheromone that aggregates other bark beetles. Other individuals join the attack, and the colonized trees have not enough resources to survive the attack (Byers, 1996). In mixed forests, it has been proved that other tree species can inhibit these pheromones, increasing the difficulty of aggregation at the same Norway spruce individual (Byers, Zhang, Schlyter, & Birgersson, 1998; Zhang & Schlyter, 2004; Zhang, Schlyter, & Anderson, 1999).

Height of Norway spruces

This predictor was also pointed by Kärvemo et al. (2014). The relation between the height of the host trees and the bark beetle infestation probability (**Figure 11b**) is caused by different reasons that have been addressed in previous studies.

The first reason relates to the bark thickness. In the research carried out by Grunwald (1986), it was shown that the occurrence of a bark beetle attack was directly related to the thickness of the bark, since bark beetles prefer trees with a relative thick bark of over 2.5 mm to create their burrows.

The height of the trees is directly related to the age of the tree (Netherer et al., 2019). Netherer & Nopp-Mayr (2005) also found that the stand age can be an accurate variable for predicting infestation probability. In this project, it was found that trees older than 60 years have a higher infestation risk. Likewise, in Seidl, Baier, Rammer, Schopf, & Lexer (2007), Bakke (1983), and Mezei et al. (2014), the older trees were more prone to be infested.

The defense mechanism that the tree possesses when an attack occurs is the segregation of monoterpene hydrocarbons in the attacked area. The amount of monoterpene that the tree can segregate depends on the resin flow. As discovered by Baier, Führer, Kirisits, & Rosner (2002), in younger tissues, the flow of the resin is better. This may be the reason why older trees are facing a higher outbreak probability.

Altitude

In **Figure 11c** a drop in the curve between approx. 720 and 850 m can be observed. This could be attributed to the microclimatic conditions of the valleys, as Overbeck & Schmidt (2012) have demonstrated, concluding that the best location for Norway spruces is shadowed moist slopes. As can be seen in **Figure 3**, these areas have altitude values within this range. This variable has also been an important predictor for Lausch, Fahse, & Heurich (2011); and Sproull, Bukowski, McNutt, Zwijacz-Kozica, & Szwagrzyk (2017). According to Mezei et al. (2014), this variable has predictive importance when the attack is in the epidemic phase.

As revealed by the correlation analysis and has been reported by Temperli, Bugmann, & Elkin, 2013, the Norway spruce biomass in the Schwarzwald National Park is higher in elevated areas. Hence, the probability of finding weakened trees is also higher. Even though the correlation analysis does not show a high correlation between altitude and sun radiation, higher areas are slightly more exposed to solar radiation. As Baier, Pennerstorfer, & Schopf (2007) suggest, within areas where sun exposure is higher, the bark temperature also increases. The solar radiation as a predictor has also been addressed by Netherer & Nopp-Mayr (2005).

Another variable that is closely related to altitude is the distance to streams. Trees located in higher elevations are further from streams (**Figure 15**), which could have an effect on the water availability for them. In addition, higher areas could be more exposed to wind. Even though a heavy storm was not registered in the study area in the previous years, trees can get damaged by low-intensity storms (Mezei, Jakuš, et al., 2017). Within the years after the damage, the tree needs to heal the wound produced by the extreme event. As shown by Baier et al., 2002, the defenses of the tree are reduced when the tree is growing.

Soil depth

In the soil depth partial dependence plot (**Figure 11d**), it can be observed that the fitted function of this variable oscillates around 0, having ranges between -0.35 and 0.2, which are very low probabilities to consider them significant. The response of this variable in the model could be influenced by the low resolution (250 m) of the input layer.

However, it has been indicated by Rehschuh, Mette, Menzel, & Buras (2017) that soil properties can negatively influence the effect of drought in Norway spruces. Namely, drought effects can be more severe in shallow sandy soils, which have a reduced water retention capacity.

Slope

The results (**Figure 11e**) demonstrated that, as found out by de Groot, Ogris, & Kobler (2018), the slope is negatively correlated with the bark beetle infestation. This result could be explained by the fact that the valleys are often narrow, being flatter slopes in higher areas with more sun radiation (**Figure 3**).

Percentage of Norway spruces

The results show that areas with a higher percentage of Norway spruces among all tree species have a higher infestation probability. In the study from Netherer & Nopp-Mayr (2005), most of the infestation areas took place where the spruce composition was above 70% of all the species.

As mentioned previously in the discussion of the variable "number of Norway spruces", a forest management reinforcing species diversity has been shown beneficial against the bark beetle infestation. However, this management is not only positive against bark beetles. Rötzer, Biber, Moser, Schäfer, & Pretzsch (2017) concluded that mixed forests also increase tree resistance against drought. The resources efficiency is higher when combining species that are complementary. The European beech is a slow-growing and shadow tolerant species. In contrast, Norway spruce is a fast-growing species that requires a lot of light (Pretzsch, 2014).

Furthermore, in mixed forests, the growth of spruces would be benefited at the beginning of the spring, when water, light, and temperature are enough for pushing the tree growth, and beeches are still leafless. Beeches could benefit from sharing the space with the spruces since the rooting system is deeper and can compete better for water resources in summer (Rötzer, Häberle, et al., 2017).

5. CONCLUSION

After following the recommendations for improving the accuracy, the model output can be used by the National Park managers. In the short term, the model can be used to increase monitoring efforts in areas that have a higher outbreak probability. As mentioned in previous sections, in the management zone of the National Park, a weekly bark beetle monitoring is carried out. The present model can reduce economic and human resources, as well as improve the response once the outbreak is in the early phases.

In the long term, the vegetation cover planning can be improved by considering the environmental factors that contribute to the higher infestation probability. The variables altitude and slope indicate that the best location for the Norway spruces is found at an elevation below 850 m and slopes above 20%. There is a body of evidence that Norway spruces have been extensively planted in Central Europe for economic purposes (Hlásny et al., 2019), suggesting that they are not native from the Schwarzwald National Park. Importantly, previous researchers addressed that areas where the Norway spruce has been artificially planted, have a higher outbreak probability (Marini et al., 2012; Ogris & Jurc, 2010). This idea is in accordance with the variables "number of Norway spruces" and "percentage of Norway spruces compared with other tree species," as well as by previous studies, which indicate that mixed-forests are more resilient (Neuner et al., 2015; Overbeck & Schmidt, 2012). This supports the objective of the Schwarzwald National Park of increasing the mixed-forest cover in development and management zones. Furthermore, the results confirm that higher trees are more prone to be infested.

It is expected that within the next years, the severity of droughts on Norway spruces in the Schwarwald National Park will increase due to its location in the warm-dry edge of the natural range (Temperli et al., 2013). In this case, the mentioned monitoring and adaptation measures are even more important. Notably, the bark beetle is not the only pest that causes problems in areas with forest plantations (European Commission, 2012; Klimo et al., 2000). This modelling approach can also be applied with the same purpose with other species.

The present study provides added insights in spatial variables that can predict bark beetles outbreaks. From a more general perspective, a structured problem has been investigated, in which different facets relating to the problem were analysed. Although these findings cannot solve the bark beetle problem in the Schwarzwald entirely, this knowledge will support the decision-making process carried out by the National Park committee.

LIST OF REFERENCES

Abdullah, H. J. (2019). *Remote sensing of European spruce (Ips typographus, L.) bark beetle green attack*. University of Twente, Enschede, The Netherlands. https://doi.org/10.3990/1.9789036547956

Akayezu, P., van Duren, I. C., Groen, T. A., Grueter, C. C., & Robbins, M. M. (2019). Abundance and spatial distribution of the main food species for mountain gorillas in the Virunga Massif, Rwanda. *Biodiversity and Conservation*, *28*(13), 3597–3620. https://doi.org/10.1007/s10531-019-01838-0

Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(3), 370–374. https://doi.org/10.1002/wics.84

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. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, *22*(1), 42–47. https://doi.org/10.1016/j.tree.2006.09.010

Araujo, M., Pearson, R., Thuiller, W., & Erhard, M. (2005). Validation of species-climate impact models under climate change. *Global Change Biology*, *11*(9), 1504–1513. https://doi.org/10.1111/j.1365-2486.2005.001000.x

Baier, P. (1996). Defence reactions of Norway spruce (Picea abies Karst.) to controlled attacks of Ips typographus (L.) (Col., Scolytidae) in relation to tree parameters, *120*, 587–593.

Baier, P., Führer, E., Kirisits, T., & Rosner, S. (2002). Defence reactions of Norway spruce against bark beetles and the associated fungus Ceratocystis polonica in secondary pure and mixed species stands. *Forest Ecology and Management*, *159*(1–2), 73–86. https://doi.org/10.1016/S0378-1127(01)00711-3

Baier, P., Pennerstorfer, J., & Schopf, A. (2007). PHENIPS-A comprehensive phenology model of Ips typographus (L.) (Col., Scolytinae) as a tool for hazard rating of bark beetle infestation. *Forest Ecology and Management*, *249*(3), 171–186. https://doi.org/10.1016/j.foreco.2007.05.020

Bakke, A. (1983). Host tree and bark beetle interaction during a mass outbreak of Ips typographus in Norway. *Zeitschrift Für Angewandte Entomologie*, *96*(1–5), 118–125. https://doi.org/10.1111/j.1439-0418.1983.tb03651.x

Bauhus, J., Forrester, D. I., Gardiner, B., Jactel, H., Vallejo, R., & Pretzsch, H. (2017). *Ecological Stability of Mixed-Species Forests*. *Mixed-Species Forests: Ecology and Management*. https://doi.org/10.1007/978-3-662-54553-9

Bentz, B. J., & Jönsson, A. M. (2015). Modeling Bark Beetle Responses to Climate Change. *Bark Beetles: Biology and Ecology of Native and Invasive Species*, 533–553. https://doi.org/10.1016/B978-0-12-417156-5.00013-7

Biedermann, P. H. W., Müller, J., Grégoire, J. C., Gruppe, A., Hagge, J., Hammerbacher, A., … Bässler, C. (2019). Bark Beetle Population Dynamics in the Anthropocene: Challenges and Solutions. *Trends in Ecology and Evolution*, *xx*(xx), 1–11. https://doi.org/10.1016/j.tree.2019.06.002

Black Forest National Park Portal. (2020).

Blanco-Canqui, H., Shapiro, C. A., Wortmann, C. S., Drijber, R. A., Mamo, M., Shaver, T. M., & Ferguson, R. B. (2013). Soil organic carbon: The value to soil properties. *Journal of Soil and Water Conservation*, *68*(5). https://doi.org/10.2489/jswc.68.5.129A

Byers, J. A. (1996). An encounter rate model of bark beetle populations searching at random for susceptible host trees. *Ecological Modelling*, *91*(1–3), 57–66. https://doi.org/10.1016/0304-3800(95)00166-2

Byers, J. A., Zhang, Q. H., Schlyter, F., & Birgersson, G. (1998). Volatiles from nonhost birch trees inhibit pheromone response in spruce bark beetles. *Naturwissenschaften*, *85*(11), 557–561. https://doi.org/10.1007/s001140050551

Caudullo, G., Tinner, W., & Rigo, D. De. (2016). Picea abies in Europe: distribution , habitat , usage and threats. *European Atlas of Forest Tree Species*, 114–116. https://doi.org/10.2788/038466

Christiansen, E., & Bakke, A. (1988). The Spruce Bark Beetle of Eurasia. *Dynamics of Forest Insect Populations*, 479–503. https://doi.org/10.1007/978-1-4899-0789-9_23

Čihák, T., & Vejpustková, M. (2018). Parameterisation of allometric equations for quantifying aboveground biomass of Norway spruce (Picea abies (L.) H. Karst.) in the Czech Republic. *Journal of Forest Science*, *64*(3), 108–117. https://doi.org/10.17221/61/2017-JFS

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., … Böhner, J. (2015). System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, *8*(7), 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015

De'ath, G. (2007). Boosted trees for ecological modeling and prediction, *88*(1), 243–251.

De'ath, G., & Katharina, F. E. (2003). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecological Applications : A Publication of the Ecological Society of America*, *13*(4), 3178–3192. https://doi.org/10.1023/A:1017558523802

de Groot, M., Diaci, J., & Ogris, N. (2019). Forest management history is an important factor in bark beetle outbreaks: Lessons for the future. *Forest Ecology and Management*, *433*(November 2018), 467–474. https://doi.org/10.1016/j.foreco.2018.11.025

de Groot, M., & Ogris, N. (2019). Short-term forecasting of bark beetle outbreaks on two economically important conifer tree species. *Forest Ecology and Management*, *450*(July), 117495. https://doi.org/10.1016/j.foreco.2019.117495

de Groot, M., Ogris, N., & Kobler, A. (2018). The effects of a large-scale ice storm event on the drivers of bark beetle outbreaks and associated management practices. *Forest Ecology and Management*, *408*(June 2017), 195–201. https://doi.org/10.1016/j.foreco.2017.10.035

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., … Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

DWD portal. Baiersbronn-Ruhestein weather station. (2020). Retrieved from https://opendata.dwd.de/climate_environment/CDC/

DWD portal. Freudenstadt weather station. (2020). Retrieved from https://opendata.dwd.de/climate_environment/CDC/

Egli, M., & Poulenard, J. (2016). Soils of Mountainous Landscapes. *International Encyclopedia of Geography: People, the Earth, Environment and Technology*, 1–10. https://doi.org/10.1002/9781118786352.wbieg0197

Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677–697. https://doi.org/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*. https://doi.org/10.1111/j.1365-2656.2008.01390.x

European Commission. (2012). *Disturbances of EU forests caused by biotic agents*. https://doi.org/citeulike-article-id:13411227

Fahse, L., & Heurich, M. (2011). Simulation and analysis of outbreaks of bark beetle infestations and their management at the stand level. *Ecological Modelling*, *222*(11), 1833–1846. https://doi.org/10.1016/j.ecolmodel.2011.03.014

FAO. (2018). Global forest resources assessment 2020. Terms and Definitions FRA 2020. Food and Agriculture Organization., 32.

Fassnacht, F. E., Latifi, H., Ghosh, A., Joshi, P. K., & Koch, B. (2014). Assessing the potential of hyperspectral imagery to map bark beetle-induced tree mortality. *Remote Sensing of Environment*, *140*, 533–548. https://doi.org/10.1016/j.rse.2013.09.014

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38–49. https://doi.org/10.1017/S0376892997000088

Freeman, E. A., & Moisen, G. G. (2008). A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, *217*(1–2), 48–58. https://doi.org/10.1016/j.ecolmodel.2008.05.015

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine, 400.

Friedman, J. H., & Meulman, J. J. (2003). Multiple additive regression trees with application in epidemiology. *Statistics in Medicine*, *22*(9), 1365–1381. https://doi.org/10.1002/sim.1501

Froeschke, J. T., & Froeschke, B. F. (2011). Spatio-temporal predictive model based on environmental factors for juvenile spotted seatrout in Texas estuaries using boosted regression trees. *Fisheries Research*, *111*(3), 131–138. https://doi.org/10.1016/j.fishres.2011.07.008

FVA Annual report. (2017).

Giunta, A. D., Jenkins, M. J., Hebertson, E. G., & Munson, A. S. (2016). Disturbance agents and their

associated effects on the health of interior Douglas-fir forests in the central Rocky Mountains. *Forests*, *7*(4), 1–30. https://doi.org/10.3390/f7040080
Greenwell, B., Boehmke, B., & Cunningham, J. (2019). Package `gbm' - Generalized Boosted Regression Models. *CRAN Repository*, 39. Retrieved from https://cran.r-project.org/web/packages/gbm/gbm.pdf%0Ahttps://github.com/gbm-developers/gbm
Grunwald, M. (1986). Ecological segregation of bark beetles (Coleoptera, Scolytidae) of spruce. *Journal of Applied Entomology*, *101*(1–5), 176–187. https://doi.org/10.1111/j.1439-0418.1986.tb00846.x
Hall, R. J., Castilla, G., White, J. C., Cooke, B. J., & Skakun, R. S. (2016). Remote sensing of forest pest damage: A review and lessons learned from a Canadian perspective *. *Canadian Entomologist*, *148*(S1), S296–S356. https://doi.org/10.4039/tce.2016.11
Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Journal of Medical Internet Research*, *20*(5), 29–36. https://doi.org/10.2196/jmir.9160
Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning. Data mining, inference, and prediction. Springer*. New York: Springer-Verlag. https://doi.org/10.1057/9780230355033.0018
Havašová, M., Ferenčík, J., & Jakuš, R. (2017). Interactions between windthrow, bark beetles and forest management in the Tatra national parks. *Forest Ecology and Management*, *391*, 349–361. https://doi.org/10.1016/j.foreco.2017.01.009
Hengl, T., De Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., … Kempen, B. (2017). *SoilGrids250m: Global gridded soil information based on machine learning*. *PLoS ONE* (Vol. 12). https://doi.org/10.1371/journal.pone.0169748
Hlásny, T., Krokene, P., & Liebhold, A. M. (2019). *Living with bark beetles : impacts , outlook and management options*.
Hobbie, S. E., Ogdahl, M., Chorover, J., Chadwick, O. A., Oleksyn, J., Zytkowiak, R., & Reich, P. B. (2007). Tree species effects on soil organic matter dynamics: The role of soil cation composition. *Ecosystems*, *10*(6), 999–1018. https://doi.org/10.1007/s10021-007-9073-4
Hofierka, J., & Súri, M. (2002). The solar radiation model for Open source GIS: implementation and applications. *Open Source GIS - GRASS Users Conference*, (September), 11–13. Retrieved from http://www.ing.unitn.it/~grass/conferences/GRASS2002/proceedings/proceedings/pdfs/Hofierka_Jaroslav.pdf
Jentschke, G., Drexhage, M., Fritz, H. W., Fritz, E., Schella, B., Lee, D. H., … Godbold, D. L. (2001). Does soil acidity reduce subsoil rooting in norway spruce (picea abies)? *Plant and Soil*, *237*(1), 91–108. https://doi.org/10.1023/A:1013305712465
Jönsson, A. M., Schroeder, L. M., Lagergren, F., Anderbrant, O., & Smith, B. (2012). Guess the impact of Ips typographus-An ecosystem modelling approach for simulating spruce bark beetle outbreaks. *Agricultural and Forest Meteorology*, *166–167*, 188–200. https://doi.org/10.1016/j.agrformet.2012.07.012
Jurc, M., Perko, M., Džeroski, S., Demšar, D., & Hrašovec, B. (2006). Spruce bark beetles (Ips typographus, Pityogenes chalcographus, Col.: Scolytidae) in the Dinaric mountain forests of Slovenia: Monitoring and modeling. *Ecological Modelling*, *194*(1-3 SPEC. ISS.), 219–226. https://doi.org/10.1016/j.ecolmodel.2005.10.014
Kärvemo, S., Van Boeckel, T. P., Gilbert, M., Grégoire, J. C., & Schroeder, M. (2014). Large-scale risk mapping of an eruptive bark beetle - Importance of forest susceptibility and beetle pressure. *Forest Ecology and Management*, *318*, 158–166. https://doi.org/10.1016/j.foreco.2014.01.025
Kautz, M., Dworschak, K., Gruppe, A., & Schopf, R. (2011). Quantifying spatio-temporal dispersion of bark beetle infestations in epidemic and non-epidemic conditions. *Forest Ecology and Management*, *262*(4), 598–608. https://doi.org/10.1016/j.foreco.2011.04.023
Klimo, E., Hager, H., & Kulhavý, J. (2000). *Spruce Monocultures in Central Europe – Problems and Prospects*. *Spruce Monocultures in Central Europe-Problems and Prospects*.
Komori, O., Eguchi, S., Saigusa, Y., Kusumoto, B., & Kubota, Y. (2020). Sampling bias correction in species distribution models by quasi-linear Poisson point process. *Ecological Informatics*, *55*(June 2019), 101015. https://doi.org/10.1016/j.ecoinf.2019.101015
Kruk, E., Ryczek, M., Klatka, S., & Malec, M. (2018). Spatial variability of chosen soil properties on mountainous area on example of the Kasińczanka stream basin, Western Carpathians. *Journal of Ecological Engineering*, *19*(3), 146–153. https://doi.org/10.12911/22998993/86154

Lausch, A., Fahse, L., & Heurich, M. (2011). Factors affecting the spatio-temporal dispersion of Ips typographus (L.) in Bavarian Forest National Park: A long-term quantitative landscape-level analysis. *Forest Ecology and Management*, *261*(2), 233–245. https://doi.org/10.1016/j.foreco.2010.10.012
Lehnert, L. W., Bässler, C., Brandl, R., Burton, P. J., & Müller, J. (2013). Conservation value of forests attacked by bark beetles: Highest number of indicator species is found in early successional stages. *Journal for Nature Conservation*, *21*(2), 97–104. https://doi.org/10.1016/j.jnc.2012.11.003
Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, *40*(4), 778–789. https://doi.org/10.1111/jbi.12058
Marini, L., Ayres, M. P., Battisti, A., & Faccoli, M. (2012). Climate affects severity and altitudinal distribution of outbreaks in an eruptive bark beetle. *Climatic Change*, *115*(2), 327–341. https://doi.org/10.1007/s10584-012-0463-z
Marini, L., Lindelöw, Å., Jönsson, A. M., Wulff, S., & Schroeder, L. M. (2013). Population dynamics of the spruce bark beetle: A long-term study. *Oikos*, *122*(12), 1768–1776. https://doi.org/10.1111/j.1600-0706.2013.00431.x
Marini, L., Økland, B., Jönsson, A. M., Bentz, B., Carroll, A., Forster, B., … Schroeder, M. (2017). Climate drivers of bark beetle outbreak dynamics in Norway spruce forests. *Ecography*, *40*(12), 1426–1435. https://doi.org/10.1111/ecog.02769
Matthews, B., Netherer, S., Katzensteiner, K., Pennerstorfer, J., Blackwell, E., Henschke, P., … Schopf, A. (2018). Transpiration deficits increase host susceptibility to bark beetle attack: Experimental observations and practical outcomes for Ips typographus hazard assessment. *Agricultural and Forest Meteorology*, *263*(June), 69–89. https://doi.org/10.1016/j.agrformet.2018.08.004
Mattivi, P., Franci, F., Lambertini, A., & Bitelli, G. (2019). TWI computation: a comparison of different open source GISs. *Open Geospatial Data, Software and Standards*, *4*(1). https://doi.org/10.1186/s40965-019-0066-y
Mezei, P., Blaženec, M., Grodzki, W., Škvarenina, J., & Jakuš, R. (2017). Influence of different forest protection strategies on spruce tree mortality during a bark beetle outbreak. *Annals of Forest Science*, *74*(4). https://doi.org/10.1007/s13595-017-0663-9
Mezei, P., Grodzki, W., Blaženec, M., Škvarenina, J., Brandyľsovaľ, V., & Jakuš, R. (2014). Host and site factors affecting tree mortality caused by the spruce bark beetle (Ips typographus) in mountainous conditions. *Forest Ecology and Management*, *331*, 196–207. https://doi.org/10.1016/j.foreco.2014.07.031
Mezei, P., Jakuš, R., Pennerstorfer, J., Havašová, M., Škvarenina, J., Ferenčík, J., … Netherer, S. (2017). Storms, temperature maxima and the Eurasian spruce bark beetle Ips typographus—An infernal trio in Norway spruce forests of the Central European High Tatra Mountains. *Agricultural and Forest Meteorology*, *242*(April), 85–95. https://doi.org/10.1016/j.agrformet.2017.04.004
Millar, C. I., & Stephenson, N. L. (2015). Temperate forest health in an era of emerging megadisturbance. *Science*, *349*(6250), 823–826. https://doi.org/10.1126/science.aaa9933
Müller, J., Bußler, H., Goßner, M., Rettelbach, T., & Duelli, P. (2008). The European spruce bark beetle Ips typographus in a national park: From pest to keystone species. *Biodiversity and Conservation*, *17*(12), 2979–3001. https://doi.org/10.1007/s10531-008-9409-1
Naimi, B., & Araújo, M. B. (2016). Sdm: A reproducible and extensible R platform for species distribution modelling. *Ecography*, *39*(4), 368–375. https://doi.org/10.1111/ecog.01881
Neteler, M., Bowman, M. H., Landa, M., & Metz, M. (2012). GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling and Software*, *31*, 124–130. https://doi.org/10.1016/j.envsoft.2011.11.014
Netherer, S., & Nopp-Mayr, U. (2005). Predisposition assessment systems (PAS) as supportive tools in forest management - Rating of site and stand-related hazards of bark beetle infestation in the High Tatra Mountains as an example for system application and verification. *Forest Ecology and Management*, *207*(1-2 SPEC. ISS.), 99–107. https://doi.org/10.1016/j.foreco.2004.10.020
Netherer, S., Panassiti, B., Pennerstorfer, J., & Matthews, B. (2019). Acute Drought Is an Important Driver of Bark Beetle Infestation in Austrian Norway Spruce Stands. *Frontiers in Forests and Global Change*, *2*(July), 1–21. https://doi.org/10.3389/ffgc.2019.00039
Neuner, S., Albrecht, A., Cullmann, D., Engels, F., Griess, V. C., Hahn, W. A., … Knoke, T. (2015). Survival of Norway spruce remains higher in mixed stands under a dryer and warmer climate. *Global Change Biology*, *21*(2), 935–946. https://doi.org/10.1111/gcb.12751
Ogris, N., & Jurc, M. (2010). Sanitary felling of Norway spruce due to spruce bark beetles in Slovenia: A

model and projections for various climate change scenarios. *Ecological Modelling*, *221*(2), 290–302. https://doi.org/10.1016/j.ecolmodel.2009.05.015

Overbeck, M., & Schmidt, M. (2012). Modelling infestation risk of Norway spruce by Ips typographus (L.) in the Lower Saxon Harz Mountains (Germany). *Forest Ecology and Management*, *266*, 115–125. https://doi.org/10.1016/j.foreco.2011.11.011

Pasztor, F., Matulla, C., Rammer, W., & Lexer, M. J. (2014). Drivers of the bark beetle disturbance regime in Alpine forests in Austria. *Forest Ecology and Management*, *318*, 349–358. https://doi.org/10.1016/j.foreco.2014.01.044

Peltzer, D. A., Bast, M. L., Wilson, S. D., & Gerry, A. K. (2000). Plant diversity and tree responses following contrasting disturbances in boreal forest. *Forest Ecology and Management*, *127*(1–3), 191–203. https://doi.org/10.1016/S0378-1127(99)00130-9

Pretzsch, H. (2014). Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. *Forest Ecology and Management*, *327*, 251–264. https://doi.org/10.1016/j.foreco.2014.04.027

Puhe, J. (2003). Growth and development of the root system of Norway spruce (Picea abies) in forest stands - A review. *Forest Ecology and Management*, *175*(1–3), 253–273. https://doi.org/10.1016/S0378-1127(02)00134-2

R Development Core Team. (2008). *R: A Language and Environment for Statistical Computing* (Vol. 2.6.2). Vienna. https://doi.org/10.1007/978-3-540-74686-7

Raffa, K. F., Aukema, B., Bentz, B. J., Carroll, A., Erbilgin, N., Herms, D. A., … Wallin, K. F. (2009). A literal use of "Forest Health" safeguards against misuse and misapplication. *Journal of Forestry*, *107*(5), 276–277. https://doi.org/10.1093/jof/107.5.276

Raffa, K. F., Grégoire, J. C., & Lindgren, B. S. (2015). *Natural History and Ecology of Bark Beetles*. *Bark Beetles: Biology and Ecology of Native and Invasive Species*. https://doi.org/10.1016/B978-0-12-417156-5.00001-0

Rehschuh, R., Mette, T., Menzel, A., & Buras, A. (2017). Soil properties affect the drought susceptibility of Norway spruce. *Dendrochronologia*, *45*(February), 81–89. https://doi.org/10.1016/j.dendro.2017.07.003

Rötzer, T., Biber, P., Moser, A., Schäfer, C., & Pretzsch, H. (2017). Stem and root diameter growth of European beech and Norway spruce under extreme drought. *Forest Ecology and Management*, *406*(September), 184–195. https://doi.org/10.1016/j.foreco.2017.09.070

Rötzer, T., Häberle, K. H., Kallenbach, C., Matyssek, R., Schütze, G., & Pretzsch, H. (2017). Tree species and size drive water consumption of beech/spruce forests - a simulation study highlighting growth under water limitation. *Plant and Soil*, *418*(1–2), 337–356. https://doi.org/10.1007/s11104-017-3306-x

Seidl, R., Baier, P., Rammer, W., Schopf, A., & Lexer, M. J. (2007). Modelling tree mortality by bark beetle infestation in Norway spruce forests. *Ecological Modelling*, *206*(3–4), 383–399. https://doi.org/10.1016/j.ecolmodel.2007.04.002

Seidl, R., Rammer, W., Jäger, D., & Lexer, M. J. (2008). Impact of bark beetle (Ips typographus L.) disturbance on timber production and carbon sequestration in different management strategies under climate change. *Forest Ecology and Management*, *256*(3), 209–220. https://doi.org/10.1016/j.foreco.2008.04.002

Seidl, R., Schelhaas, M. J., & Lexer, M. J. (2011). Unraveling the drivers of intensifying forest disturbance regimes in Europe. *Global Change Biology*, *17*(9), 2842–2852. https://doi.org/10.1111/j.1365-2486.2011.02452.x

Seidl, R., Schelhaas, M. J., Lindner, M., & Lexer, M. J. (2009). Modelling bark beetle disturbances in a large scale forest scenario model to assess climate change impacts and evaluate adaptive management strategies. *Regional Environmental Change*, *9*(2), 101–119. https://doi.org/10.1007/s10113-008-0068-2

Sproull, G. J., Bukowski, M., McNutt, N., Zwijacz-Kozica, T., & Szwagrzyk, J. (2017). Landscape-level spruce mortality patterns and topographic forecasters of bark beetle outbreaks in managed and unmanaged forests of the Tatra Mountains. *Polish Journal of Ecology*, *65*(1), 24–37. https://doi.org/10.3161/15052249PJE2017.65.1.003

Stadelmann, G., Bugmann, H., Meier, F., Wermelinger, B., & Bigler, C. (2013). Effects of salvage logging and sanitation felling on bark beetle (Ips typographus L.) infestations. *Forest Ecology and Management*, *305*, 273–281. https://doi.org/10.1016/j.foreco.2013.06.003

Stadelmann, G., Bugmann, H., Wermelinger, B., & Bigler, C. (2014). Spatial interactions between storm

damage and subsequent infestations by the European spruce bark beetle. *Forest Ecology and Management*, *318*, 167–174. https://doi.org/10.1016/j.foreco.2014.01.022
Stadelmann, G., Bugmann, H., Wermelinger, B., Meier, F., & Bigler, C. (2013). A predictive framework to assess spatio-temporal variability of infestations by the european spruce bark beetle. *Ecography*, *36*(11), 1208–1217. https://doi.org/10.1111/j.1600-0587.2013.00177.x
Stängle, S. M., Sauter, U. H., & Dormann, C. F. (2017). Comparison of models for estimating bark thickness of Picea abies in southwest Germany: the role of tree, stand, and environmental factors. *Annals of Forest Science*, *74*(1). https://doi.org/10.1007/s13595-016-0601-2
Stephanie, M., Ceri, W. H., & S.J., O. (2001). Evaluating presence – absence models in ecology : the need to account for prevalence. *Journal of Applied Ecology*, *38*, 921–931. https://doi.org/10.1080/09613210110101185
Temperli, C., Bugmann, H., & Elkin, C. (2013). Cross-scale interactions among bark beetles, climate change, and wind disturbances: A landscape modeling approach. *Ecological Monographs*, *83*(3), 383–402. https://doi.org/10.1890/12-1503.1
Thom, D., Seidl, R., Steyrer, G., Krehan, H., & Formayer, H. (2013). Slow and fast drivers of the natural disturbance regime in Central European forest ecosystems. *Forest Ecology and Management*, *307*, 293–302. https://doi.org/10.1016/j.foreco.2013.07.017
Veloz, S. D. (2009). Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, *36*(12), 2290–2299. https://doi.org/10.1111/j.1365-2699.2009.02174.x
Veran, S., Piry, S., Ternois, V., Meynard, C. N., Facon, B., & Estoup, A. (2016). Modeling spatial expansion of invasive alien species: Relative contributions of environmental and anthropogenic factors to the spreading of the harlequin ladybird in France. *Ecography*, *39*(7), 665–675. https://doi.org/10.1111/ecog.01389
Vorster, A. G., Evangelista, P. H., Stohlgren, T. J., Kumar, S., Rhoades, C. C., Hubbard, R. M., … Elder, K. (2017). Severity of a mountain pine beetle outbreak across a range of stand conditions in Fraser Experimental Forest, Colorado, United States. *Forest Ecology and Management*, *389*, 116–126. https://doi.org/10.1016/j.foreco.2016.12.021
Waldmanagement im Nationalpark Schwarzwald. (2017).
Wegensteiner, R., Wermelinger, B., & Herrmann, M. (2015). *Natural Enemies of Bark Beetles: Predators, Parasitoids, Pathogens, and Nematodes. Bark Beetles: Biology and Ecology of Native and Invasive Species*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-417156-5.00007-1
Wei, T. (2017). Title Visualization of a Correlation Matrix. *Statistician*, *56*, 316–324. Retrieved from https://github.com/taiyun/corrplot%0Ahttps://github.com/taiyun/corrplot/issues
Wickham, H., Chang, W., Henry, L., Lin Pedersen, T., & Takahashi, K. (2020). Package 'ggplot2.'
Zhang, Q. H., & Schlyter, F. (2004). Olfactory recognition and behavioural avoidance of angiosperm nonhost volatiles by conifer-inhabiting bark beetles. *Agricultural and Forest Entomology*, *6*(1), 1–20. https://doi.org/10.1111/j.1461-9555.2004.00202.x
Zhang, Q. H., Schlyter, F., & Anderson, P. (1999). Green leaf volatiles interrupt pheromone response of spruce bark beetle, Ips typographus. *Journal of Chemical Ecology*, *25*(12), 2847–2861. https://doi.org/10.1023/A:1020816011131
Zýval, V., Křenová, Z., & Kindlmann, P. (2016). Conservation implications of forest changes caused by bark beetle management in the Šumava National Park. *Biological Conservation*, *204*, 394–402. https://doi.org/10.1016/j.biocon.2016.11.001

6. APPENDIX

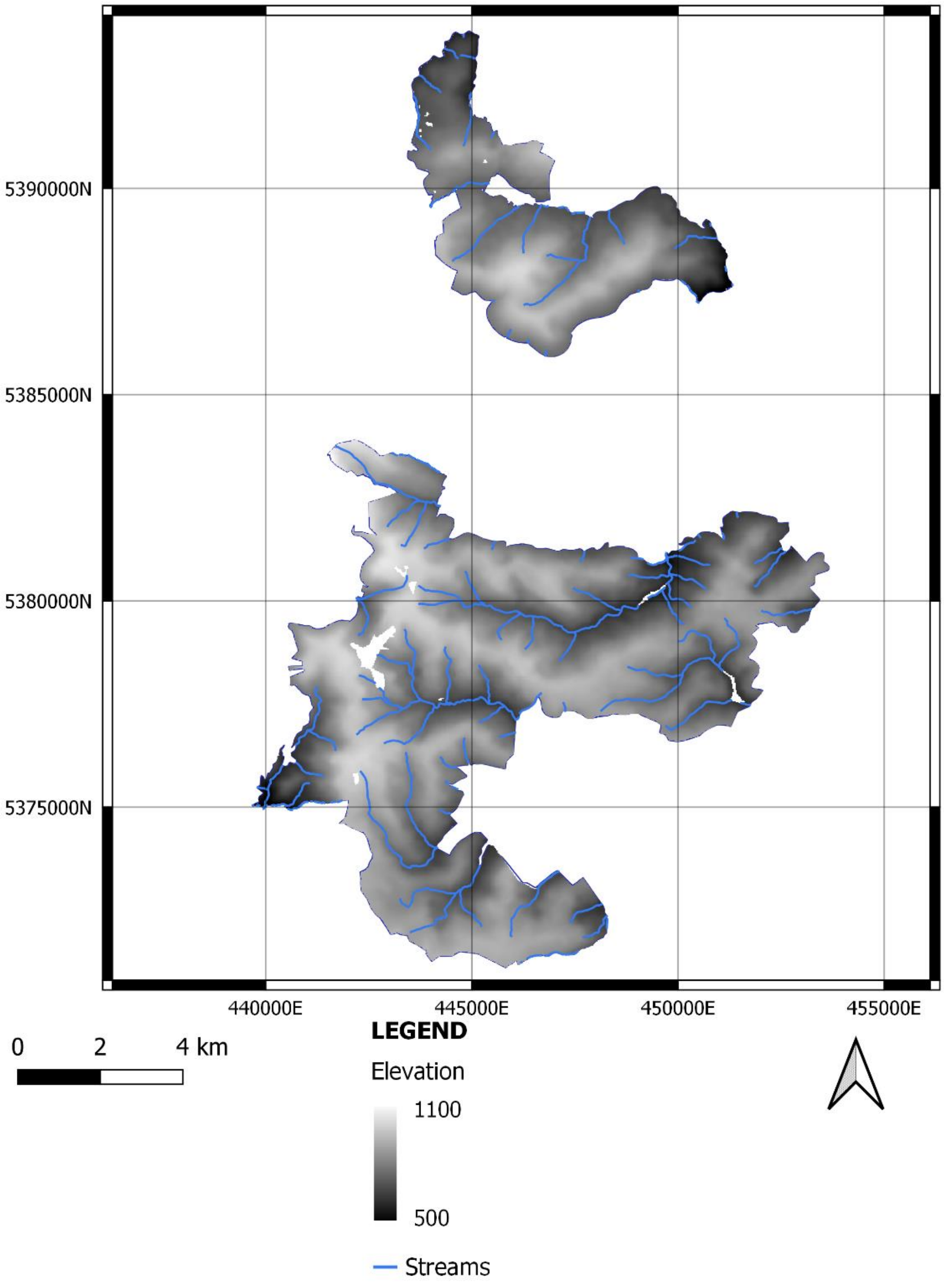


Figure 15. Map of streams on the digital elevation model in the Schwarzwald National Park.