The effects of land use, wildfire and honeybee colonies in the distribution of *Vespa velutina* in Portugal

Tomás Duarte Santiago

Dissertação para a obtenção do Grau de Mestre em
Gestão e Conservação de Recursos Naturais

Orientadores: Professora Manuela Branco (ISA)
Dra. Anabela Nave (INIAV)

Júri:
Presidente: Doutor José Maria Horta e Costa Silva Santos, Professor auxiliar do(a) Instituto Superior de Agronomia da Universidade de Lisboa
Vogais: Doutora Manuela Rodrigues Branco Simões, Professora associada com agregação do(a) Instituto Superior de Agronomia da Universidade de Lisboa, orientadora;
Doutor Professor Pedro Segurado, Professor auxiliar do(a) Instituto Superior de Agronomia da Universidade de Lisboa

2024

Acknowledgments

The completion of this work was by far the toughest academic challenge I’ve faced. Alone it would’ve never happened. With that in mind I’d like to start by expressing my gratitude to Professor Manuela Branco and Dra. Anabela Nave, it has been an honour and my greatest pleasure to work alongside the two of you. Thank you for always checking in on me when I wouldn’t email for weeks. Thank you for believing in me and providing the much-needed confidence in me.

I’m also incredibly grateful for all the assistance Professor Rui Figueira provided me when curating the hornet data. Likewise, I’d like to extend a special thanks to Professor Pedro Segurado for assisting me with the geographical analysis and statistical modelling, introducing me to a world I thought I’d never understand.

I’d also like to extend a huge thanks to all the institutions that provided the data used for this work. Especially the team at INIAV for helping me obtain both the honeybee data as well as the hornet sightings. As well as my team at LandCOBA for distracting me from this madness from time to time and my boss for allowing me to work on this during company time.

It goes without saying that without my parents none of this would be possible. Thank you to the both of you for providing me with the opportunities, the self-confidence and reassurance that no matter what I chose to do with my life, they’d always support me.

To my loving girlfriend who listened to me ramble about hornets and curse like a sailor whenever my models would spit out an error message, and of course, our cat Bee for always sitting on my computer, I thank you endlessly for all the support you’ve given me throughout this journey.

My masters’ colleagues, thank you all for embarking on this two-year journey with me. I’ll cherish all the memories we’ve made, and I’ll keep them with me forever.

Finally, I’d like to thank all my friends at BUS for providing me with endless afternoons to ramble about hornets and every step of my process.

Thank you all for making all this possible!!

Resumo

Vespa velutina nigrothorax é uma espécie invasora que encontra-se estabelecida na Península Ibérica desde 2011. Um dos seus impactos negativos mais notáveis, a nível ecológico e económico, resulta da predação das populações de abelhas (*Apis mellifera*). Contudo, há um entendimento limitado sobre o papel que esta espécie desempenha no declínio das populações de abelhas. Apesar de existirem estudos que examinam os fatores ambientais que contribuem para a invasão da vespa-asiática, os efeitos do uso do solo a uma escala mais fina continuam pouco compreendidos.

Este estudo tem como objetivo analisar como o uso do solo e outros fatores como os incêndios florestais e densidade de colónias de abelhas moldaram a distribuição da *V. velutina* e os seus padrões a nível municipal. Pretende-se também estudar a localização dos ninhos dentro do raio de atividade da vespa (aproximadamente 500m de diâmetro) para compreender os padrões de uso do solo dentro deste perímetro. Por fim, busca-se compreender a influência da vespa nas populações apícolas locais.

Para este fim, foram utilizados dados recolhidos por ciência-cidadã relativos à presença das vespas, em conjunto com variáveis explicativas relacionadas com o uso do solo e outros fatores ambientais.

As “florestas de coníferas” foram identificadas como tendo o maior impacto negativo na abundância de ninhos de vespas. Por outro lado, os incêndios florestais não parecem ter uma influência significativa. A densidade populacional humana foi identificada como um fator determinante importante. Na vizinhança imediata dos ninhos, as vespas demonstram uma forte preferência por “áreas artificiais” e “terrenos agrícolas” revelando uma forte relação com áreas de maior ocupação humana.

Finalmente, as populações de abelhas demonstraram um papel significativo na abundância desta vespa. Por outro lado, revelou-se uma relação entre o declínio das colónias de abelhas e o aumento da população de vespas, embora este seja apenas marginalmente significativa.

Palavras-chave – Apiários, abelhas, declínio, vespa, espécie invasora.

Abstract

The invasive species, *Vespa velutina nigrothorax*, has been established in the Iberian Peninsula since 2011. One of its most notable ecological and economic negative impacts results from its predation on honeybees (*Apis mellifera*). Despite this, there is limited understanding of the role of this invasive species plays on the decline of western honeybees' populations. Moreover, despite numerous studies examining the environmental factors contributing to the invasion of the Yellow Legged Hornet, the effects of land use at a finer scale remain poorly understood.

This study aims to analyse how land use and wildfires have shaped the distribution of *V. velutina* at a municipality level. Also, study nest's location within the hornet´s activity range (500m diameter) to gain an understanding on the land use and wildfire patterns within this buffer. Finally, to understand the influence of the hornet on local apiary populations and the significance of these apiaries in its dispersal.
For this, crowd-sourced data on hornets, in tandem with explanatory variables related to land use and other environmental factors were used. We anticipate shedding light on observed patterns and gaining a better understanding of the intricate dynamics involving the YLH.

"Conifer forests" were identified as having the greatest negative impact for the sighting of nests while "Wildfires" did not seem to have a significant influence. "Human population" density was also determined to be a significant driver. On the immediate vicinity of the nests the hornets appear to preferer "Artificial areas" as well as "Agricultural land" revealing a strong affinity to areas with human settlements.

Finally, honeybee populations appear to be strongly linked to the abundance of YLH nests. On the other hand, a relationship between the decline of honeybee colonies and the rise in hornet population was shown to be marginally significant.

Key words – Apiaries, bees, decline, hornet, invasive species.

Resumo alargado

As espécies animais que transpõem fronteiras naturais e acabam por prosperar num novo ambiente podem ter impactes na biodiversidade local (que na maioria dos casos o efeito são negativos). Quando uma espécie tem predominantemente efeitos negativos não negligenciáveis e continua a invadir novas áreas passa a ser considerada espécie invasora. Os insetos são candidatos perfeitos para se tornarem espécies invasoras. Dada a sua pequena dimensão passam despercebidos facilitando a sua introdução em novas zonas quer pela ação do Homem, tais como em trocas comerciais, quer pela sua dispersão natural. Um exemplo recente é o da espécie *Vespa velutina nigrothorax*, considerada uma espécie invasora na Europa desde 2004 e na Península Ibérica desde 2011. Introduzida no continente europeu através de um carregamento de mercadorias vindas da China para França, esta espécie em pouco tempo se adaptou às condições locais. Qualquer espécie invasora prejudica o ecossistema que invade de diversas maneiras. Competição por recursos, predação, disseminação de pragas e/ou doenças, os impactos negativos que causa no local invadido são características importantes ter em conta. No caso da *V. velutina*, um dos impactos ecológicos e económicos mais relevantes é o da sua predação sobre insetos polinizadores, particularmente a abelha-europeia *Apis mellifera*. Enquanto que a abelha melífera asiática (*Apis cerana*) evoluiu com a presença da vespa-asiatica e desenvolveu medidas contra os seus ataques, a abelha-europeia nunca esteve na presença de um predador tão eficaz como a *V. velutina*. Essa falta de defesas deixou a abelha europeia vulnerável à predação excessiva pela *V. velutina* levando ao colapso de colónias em casos mais graves. Quando essa predação leva ao colapso das colónias, pode levar também ao abandono da prática da apicultura local agravando ainda mais o défice de polinizadores a nível regional e nacional. Apesar de presente em Portugal há mais de uma década, há um conhecimento limitado do papel que esta espécie invasora desempenha no declino das populações de abelha-do-mel e de como as práticas apícolas podem estar a contribuir para a disseminação desta espécie. Visto que a primeira *V. velutina* foi detetada há mais de uma década em Viana do Castelo, o efeito que esta tem na apicultura já pode ser sentido. Já se sabe como fatores climáticos influenciam a *V. velutina*, tendo uma preferência por climas mais húmidos e temperaturas mais amenas, ou como a altitude limita a sua dispersão, tendo preferência por zonas de quotas mais baixas. Contudo, pouco é conhecido sobre como os diferentes usos de solo, população humana e incêndios rurais contribuem e condicionam a invasão, contributo que este estudo pretende dar, analisando como estes fatores impactam os padrões de dispersão à escala municipal.

Para além disso, a localização dos ninhos e a paisagem em que se inserem foi pouco estudado. Estudos prévios que utilizaram a mesma base de dados que este estudo, já concluíram que a *V. velutina* tem preferência por nidificar nas copas das árvores uma vez que estas fornecem abrigo e camuflagem aos ninhos. Estruturas antropogénicas como muros, telhados e interiores de imóveis também foram identificadas como sendo estruturas de preferência para nidificar. Segundo uma base de dados mais atual, a quantidade de ninhos detetados nas copas das árvores diminuiu ligeiramente enquanto aquele encontrado em estruturas antropogénicas aumentou. Com o objetivo de perceber qual a influência da paisagem na localização dos ninhos validados, foi feito um buffer de 500m de diâmetro à volta de cada ninho, correspondente ao raio de atividade principal da vespa -asiática, tendo sido analisado uso do solo bem como a incidência de incêndios nesse buffer. Finalmente, também foi estudada a influência da vespa-asiática nas populações locais da abelha-do-mel e nos apiários em que se inserem, com o objetivo de perceber qual o panorama atual e como é que este foi influenciado pela vespa-asiática. Ao mesmo tempo também foi estudada a influência dos incêndios rurais em ambas as espécies (*A. mellifera* e *V. velutina*).

Cada um dos estudos mencionados foi desenvolvido em áreas de estudo semelhantes compreendidas em seis distritos de Portugal continental, três no centro oeste e três no Noroeste. A localização dos ninhos e a influência da paisagem na disseminação da vespa-asiática foi estudada para os distritos de Braga, Porto, Viana do Castelo, Leiria, Coimbra e Santarém. Para a influência da paisagem, no distrito de Santarém, usou-se só a zona a norte do rio Tejo uma vez que a vespa-asiática tem dificuldade em transpor o rio e podia levar a um enviesamento de avistamentos falsos nessa zona. Por último, a influência da vespa-asiática na apicultura e nas abelhas foi estudada nos três distritos do noroeste, Porto, Braga e Viana do Castelo, distritos com o maior período de invasão da espécie. A escolha de duas áreas de estudo em localidades diferentes teve como objetivo de analisar paisagens diferentes a nível do uso do solo bem como densidades populacionais e de incêndios distintas de modo a possibilitar compreensão mais detalhada dos possíveis efeitos. Os dados de avistamentos de ninhos de vespas-asiática foram obtidos dos registos na plataforma Stopvespa, https://stopvespa.icnf.pt/, plataforma de ciência cidadã, passando por uma curadoria rigorosa e analisados com o uso de Sistemas de Informação Geográfica (SIG), incorporando variáveis explicativas relacionadas com o uso do solo e de outros fatores ambientais. Através deste estudo, pretendeu-se esclarecer certos padrões e obter uma melhor compreensão das dinâmicas complexas que envolvem a *V. velutina*, contribuindo assim para estratégias de gestão mais eficazes.

As, "florestas de coníferas" foram identificadas como sendo a variável de uso do solo mais influente, embora com um efeito negativo, na abundância de ninhos de vespas-asiáticas. Possivelmente a reduzida biodiversidade encontrada nestas florestas dificulta que a espécie obtenha condições para as suas necessidades alimentares. Por outro lado, os incêndios rurais não aparentam ter qualquer tipo de impacto na dinâmica da espécie. As vespa-asiáticas tendem a ocorrer em maior abundância em locais com maior densidade populacional humana e com maior proporção de áreas artificiais, como casas habitacionais, indústrias e equipamentos culturais ou de laser, a nível municipal. Isso pode refletir uma preferência da espécie ou simplesmente um enviesamento introduzido pelo uso de dados de ciência cidadã. O estudo da localidade dos ninhos também veio refletir a conclusão de que as vespas-asiáticas tendem a nidificar em estruturas próximas de áreas artificiais como habitações e terrenos agrícolas (a que se associa biodiversidade mais alta quando comparada com florestas puras ou habitações) enquanto que, a combinação de "culturas permanentes" com "florestas de coníferas" foi identificada como desfavorável à vespa-asiática, possivelmente devido à pouca biodiversidade a que se pode associar a povoamentos puros de resinosas assim como pomares de compassos mais apertados.

Por fim, as populações de abelhas também parecem desempenhar um papel significativo positivo na abundância da vespa-asiática, uma vez que a taxa de avistamento de ninhos aumenta com a quantidade de colónias e apiários nos municípios. Os apiários acabam por ser zonas com uma densidade elevada de abelhas que não exigem muito esforço por parte das vespas-asiáticas para serem caçadas.

Os dados refletem um decréscimo nas populações de colónias de abelhas na zona Norte, mas não no declínio de apiários. O declínio no número de colónias pode ser em parte justificado pela predação por parte da vespa-asiática. Por último, o efeito dos incêndios não se pode associar às variações que se observam quer nas populações de abelhas como nos ninhos de vespas-asiáticas.

Table of Contents

Table index

Figure index

1. Introduction

Since humans started traveling throughout the globe and from continent-to-continent, animal species have been traveling along with us. Sometimes the transport of these animals is intentional while in most cases its unintentional. In many cases, this has led to invasive processes, with negative ecological and economic impacts being observed (Kenis et al., 2007). This is a direct consequence to globalization and the trades of goods and services between continents. Either through shipping containers or flights, plants, animals, fungi, bacteria and viruses now have an assisted route to colonize places where they have not been before. The introduction and trade of these species between continents dates as far back as the 1500, when European explorers would take mammals and birds on their boats to the new world (Drake, 1989). However, for the last five decades the invasion of alien species has intensified across the world (Seebens et al., 2017) due to globalisation trends and international movement of people and goods. In the European continent, most insects are brought as stowaways in the ornamental plant trade. Of those insects that invade Europe on a yearly basis, around 40% originate from Asia (Kenis et al., 2007). To appropriately manage and respond to these invasions we, require a solid understanding of the drivers that influence them. From the establishment process and spread, to the intrinsic traits of a species, all these factors are fundamental to control the invasion.

One of the recently established invasive species in Europe is the Yellow Legged Hornet (YLH), *Vespa velutina Lepelletier* (Hymenoptera, Vespidae). First introduced to the European continent through France, around 2004, it is believed this hornet species was brought on a shipment of vases coming from China (Villemant et al., 2006) possibly Zhejiang or Jiangsu province (Monceau et al., 2014). Currently the YLH has been sighted in nine European countries (López et al., 2011, Darrouzet et al., 2015).

In 2011 the hornet was first observed in continental Portugal in the Northern District of Viana do Castelo (Carvalho et al., 2020). In 2024 the YLH has already migrated South of the Tagus River. Several studies have been carried out to understand both the biology and the invasion patterns of the YLH. Some studies that have already been carried out in Portugal have analysed the distribution and invasion patterns using climatic drivers (Catarino, 2022), while others study the selection of nesting sites in a smaller scale (Carvalho et al., 2020). Still, there is a lack of knowledge on the effects that these invasive animals have had on local honeybee populations as well as the effect wildfires have had on the YLH. Finally, the interactions between the invasive YLHs and land use (LU) have not yet been studied in a municipality scale.

1.1. Biology and identification of the Yellow Legged Hornet

Currently in Portugal there exists two species of hornets, the invasive *V. velutina nigrithorax* commonly referred to as the Yellow Legged Hornet, and the native *Vespa crabro* L. known by the common name European hornet. They are easily distinguishable from one another due to their size and colouration difference. The native European Hornet being slightly larger (3 – 3.5cm) with darker legs and a yellower abdomen, alternatively, the *V. velutina* is slightly shorter (2.5 – 3cm) with a black colouration to its thorax, the last third of its abdomen and its legs being yellow (Verdasca, 2022).

When it comes to nesting, *V. velutina* will preferer tree canopies, with around 70% of reported nests being found there. A further 20% of reported nests being found in or related to anthropogenic structures such as homes, buildings or walls. The remainder found either on bushes or other shrub like vegetation and in the ground on the base of trees and other structures (Carvalho et al., 2020). The primary nest is built in the beginning of the year between the months of February and May by the foundress, it tends to be smaller in size at around 10cm in a protected place. This first nest has the function of providing a safe place for the foundress to lay its first generation of eggs leading to the first generation of workers. These workers will help grow the colony until it is time to build the secondary nest between the summer months (Verdasca, 2022). This secondary nest tends to be built higher above the ground, usually around 10m high, in tree canopies and other shielded areas (Carvalho et al. 2020). Nests found around the months of August to September were around 40-50cm in height with a diameter ranging between 25-30cm in length with 4 to 5 combs within it. Whereas nests found in the end of November had reached an average height of 60-90cm and a radius between 40-70cm with 6 to 7 combs inside (Choi et al., 2012).

Vespa velutina nigrothorax falls into the category of social Hymenoptera having a clear distinction between the workers (infertile females), males and the queen (fertile female). A mature colony is composed on average by one queen, around 900 males and 1000 to 1200 workers (Choi et al., 2012). The distinction between a male and female YLH are found in the existence of a stinger in the females and longer antenna in the males. Coloration and size don’t differ between genders. Between the workers and the queen, little to no differences can be found throughout most of the year. However, inside the colony the queen can be identified by their slightly distended abdomen and damaged wings at the tail end of the season (Bunker, 2022).

The biological cycle of *V. velutina* varies according to climate. In a tropical climate, YLH colonies are found in all stages of development all over the year. Whereas in a temperate region, the YLH colonies will follow an annual cycle (Verdasca, 2022). The cycle begins with the emergence of the founder queen from hibernation. *V. velutina* queens hibernate seemingly anywhere as long as it is small, dark and secluded from potential predators (Franklin et al., 2017). Once she has awoken, she will begin the construction of the primary nest and the first brood of workers emerging between the months of February and May. Between the months of April and September there is the construction of the secondary nest, and this is the period where the colony grows the most. Between the months of August to November is when the colony matures, and the first sexually viable adults appear. Next, between the months of September and November the mating process takes place where the queen's will be mated and will be finally capable of founding a new colony in the upcoming year. Finally, the death of the colony will occur between November and January where the mated queens will enter hibernation awaiting the warm weather from the next year to start colonizing again (Bunker, 2022).

1.2. Distribution and invasion

The native range of the YLH goes as far West as Afghanistan and as far East as the islands of Java, the Lesser Sunda Islands, Sulawesi and Sumatra.

Meanwhile, in the Asian continent, it is considered invasive in both Japan (Kishi & Goka, 2017) and the Korean Peninsula (Choi et al., 2012, Arca et al., 2015). In the European continent the YLH can be found as far West as Portugal (Carvalho et al., 2020) and as Far East as Czech Republic (Walter et al., 2024) as well as the United Kingdom (Budge et al., 2017).

The fact that designated queens tend to hibernate during the months of winter makes them a suitable candidate for long distance dispersion. For example, a hibernating queen can spend long periods of times in shipping containers on either ships or trucks making it easy to traverse long distances (Verdasca, 2022). Arca et al. (2015) has suggested that as little as a single, multi-mated queen was responsible for the invasion of the YLH in mainland France. From this one mated queen, a colony can arise with over 13 000 individuals (Rome et al., 2015) making this particularly prolific invasive alien species.

1.3. Impacts of the Yellow Legged Hornet

Like most hornets, the adults of this species will primarily feed on sugars derived from flower nectars, tree saps and fruits but they will prey on other Hymenoptera, mainly honeybees, as a source of protein for the larvae (Beggs et al., 2011). This species is a particularly proficient hunter as it has been documented being faster and more agile than other hornets (Verdasca, 2022) and will hover over the entrance of a hive and snatch honeybees mid-flight in an act of bee-hawking (Tan et al., 2007). As such, multiple publications have attributed the YLH to the steady decline of honeybees in a particular area not only from direct hunting, but also from impeding the bees from foraging, reducing the availability of resources for the colony during the winter months (Monceau & Thiéry, 2017). With this, comes all the related impacts on local honey production as well as local pollinator populations. If the problem persists and intensifies over the years this could have a substantial impact on local, agricultural driven economies.

As for the impacts on public health, *V. velutina* has not shown much evidence of being more of a threat to humans than the native *V. crabro* (Quintans, 2020). However, the danger arises when taken into consideration the fact that *V. velutina* tends to be more aggressive than *V. crabro*. The amount of venom it injects per sting and the fact that its stinger is slightly longer than others make it so a larger quantity of venom is injected into the (Galvão 2021). Also, the colonies of *V. velutina* are larger than *V. crabo*, increasing the odds of multiple stings.

1.4. Modelling the distribution and spread of the Asian hornet

Several studies aimed at modelling the distribution of the YLH have been done both in Europe and Asia. Bessa et al. (2015) modelled the distribution for the YLH in the Iberian Peninsula using a generalised linear model and concluded that the northern areas of the peninsula provide a suitable habitat for the expansion of the species. Meanwhile, the dryer southern regions of Portugal and Spain appear to be spared from the invasion for now. Climatic drivers such as precipitation and summer harsh temperature seem to have the highest influence in the distribution (Catarino, 2022).

Another study conducted by Requier et al. (2019) concluded, using modelling tools, that the predation of the YLH on western European honeybees and the consequent foraging paralysis behaviour is an undoubtable mechanism that leads to the collapse of the colonies in the winter months. Through field observations they were able to conclude that predator-prey relationships near the entrance of the hives led to homing failures from the bees and eventual foraging paralysis.

1.5. Objectives and motivation

Despite its occurrence in the European continent for over a decade, being present in over half of continental Portugal (Carvalho et al., 2020), few studies have delved deep into the preferences of land use, for nesting. Catarino, (2022) modelled the distribution of the YLH in continental Portugal for the districts of Coimbra and Viana do Castelo using climatic and altitudinal variables and concluded that high elevations weren't suitable for the YLH. Furthermore, they concluded that the precipitation in the humid months had a positive relation to the presence of the YLH. Meanwhile, Galvão (2021), studied possible action plans in the mitigation and possible eradication of *V. velutina* in Portugal.

As it stands, there hasn't been a study that deals with the land uses around the activity radius of a nest as well as the land use preference when choosing a nest. While Catarino, (2022) has used human population data, it was only for two urban centres in Portugal.
With this study, we aim to shed a light on:

1) The local distribution of the YLH nests in relation to land use patterns and wildfire occurrences.
2) The potential YLH impact on honeybee populations
3) How land use, human population, wildfire and honeybee populations affects the density of the nest (number per km^2) sightings at a municipality scale.

2. Materials and methods

2.1. Study Area

This study is comprised of three different studies using the same data sources. “Hornet nest density” was carried out in all municipalities from Porto, Braga, Viana do Castelo, Leiria and Coimbra as well as those municipalities north of the Taugs river as it is believed that it may serve as a physical barrier for the hornet. The “Local position of the nests” was carried out in 96 municipalities from six districts, Braga, Porto, Viana do Castelo, Coimbra, Leiria and Santarém. The first three pertaining to the Northen region of the country, where the YLH was first sighted, and the latter three corresponding to the centre region of the country. The third study, “Trends in honeybee populations”, will only use information on the three northern districts in which in YLH is present for a longer time, that was assumed to be long enough for having an impact.

The selection of three northern and three central districts was based on the diverse land uses in these two distinct regions of the country. This selection provides a broad range of land uses for the study and ensures adequate representation of certain classes that may have a regional bias. The north is characterized by a heavier emphasis on small scale, temporary agriculture, while the centre is characterized by a predominant permanent agriculture dominated by orchards. When it comes to forestry, there is some difference between the two regions with the centre boasting higher proportions of forests or forest tree monocultures with the predominant type being “Hardwood forests”. Despite this, the predominant species is *Pinus pinaster* with over 179 000ha according to the 2015 national forestry inventory (IFN2015).

When studying the effects of the YLH on the local population of domesticated honeybees it was important to focus on an area where the YLH has been present the longest. For that effect, the three northern districts of Viana do Castelo, Porto and Braga were chosen. These three districts have reported nest sightings as far back as 2015 for Viana do Castelo and Porto and 2014 for Braga. This is assumed to be enough time for the YLH to establish a stable population.

For the hornet nest density analysis, the district of Santarem was cut following the path of the Tagus River. This was done because the river seems to function as a barrier. Although the YLH has already trespassed the river, the number of reported nests is low. This would lead to a bias when it came to projecting the pseudo-absences. As such, the decision was made to only use the area north of the Tagus River for the district of Santarem. The following map illustrates the three study areas (Fig. 1).

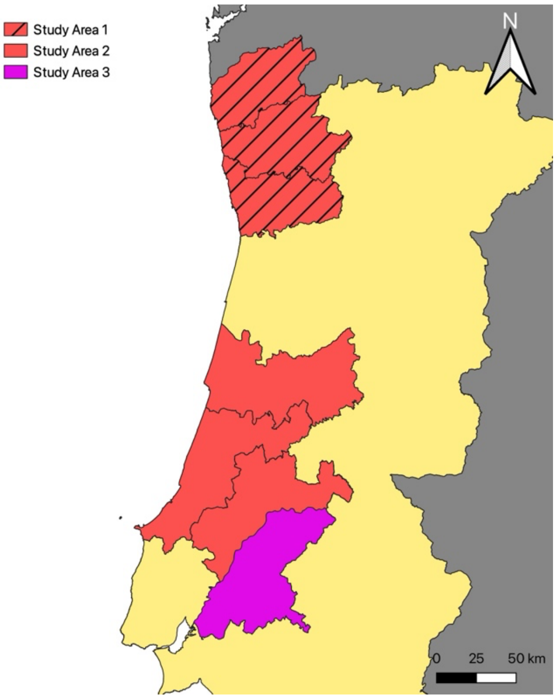


Figure 1 - Study area. Trends in honeybee populations – Area 1; Hornet nest density – Areas 1 and 2; Local position of the nests – Areas 1, 2 and 3

2.2. Data source

2.2.1. Yellow Legged Hornet

The base of this study, the YLH sighting data, were sourced from the STOPVespa initiative overseen by the National Institute for the Conservation of Forests (ICNF). This is a citizen science program where individuals are encouraged to report sightings of either individual hornets or hornet nests they may encounter during their day. This spreadsheet contained all data recorded by the STOPVespa initiative since its inception in 2013 up to 2023 by ICNF. This spreadsheet contains the nature of each YLH sighting by describing the location (in geographic coordinates WGS84) of the nest, the type of nest (primary or secondary), the dimensions of each nest, the location in which they were found (tree, rooftop, wall, etc..) and, whether the nest was destroyed or not. Nests that weren’t destroyed means they were confirmed not to be *V. velutina* nests and therefore were left alone.

2.2.2. Honeybees

The second most important data set used was the national statistics on the number of honeybee colonies, apiaries and hives in each parish for Portugal, available since 2015. This data was provided by the General Directorate of Food and Veterinary (DGAV) in the form of spreadsheets and solely contains the absolute value for each of the categories mentioned previously for each parish as well as the DICOFRE (numeric code unique to each parish) in Portugal, this data comprises the period between the years of 2013 and 2023. No major curation process was required. The only curation done was selecting all the parishes belonging to each municipality and adding their values together to get a total for honeybee apiaries and colonies for each of the 96 Municipalities.

Finally, the last data set in the form of spreadsheet used was the national census data collected by the National Institute of Statistics (INE) for the year of 2019. Comprised fully by the absolute value for the human population of each municipality for continental Portugal.

2.2.3. Landscape data

The following data sets used are GIS data of open access provided by the Portuguese state. First and foremost, the land use/cover assessment chart of COS2018.v2 for the country was provided in the form of a GeoPackage by the DGT. Each polygon was classified using a hierarchical classification with 4 levels of specification and a minimum area of 1ha.

Next was the wildfire incidence for continental Portugal from 2019 to 2023 available at the ICNF geocatalogue. Each year was merged into one and subsequently clipped to each relevant district

Finally, the borders for each district and municipality in Portugal, can be found the official administrative chart for the country (CAOP2022) published in 2022 by DGT.

2.3. Data Curating

Two individual data sets were used to create the final set of curated hornet nest sightings. The first one was compiled in 2023 and after some curation it was concluded that it was incomplete. Some municipalities only uploaded data for 2-3 years and the remainder was left blank. When separating the sightings by year some gaps started to become apparent. As such, a second request was made to ICNF in 2024 for the municipalities to submit the sightings. The second data set was also incomplete and had gaps in specific years for specific municipalities that did not correspond to the gaps in the first data set. As such, both were combined into one and a curation process was initiated.

Both tables contain the same type of information pertaining to sightings of YLH nests as well as individual hornets. With a variable number of attributes per table, some had to be given priority over others. From this, only fields related to the nature of the sighting (date, location and size) as well as to the nature of the destruction (date and whether they were destroyed) as well as geographical fields (coordinates and district of sighting). Lastly, id fields such as global ids and parent ids were kept as an easy way to identify each sighting in case some inconsistency was found.

Given that this data is derived from citizen science and relies on observations made by individuals, there can be issues. The main problem is the potential misidentification of the nests. Consequently, only nests that were destroyed and thus confirmed by authorities as *V. velutina* were used. The next issue is related to the method used to collect this data. Relying on a survey form that is filled out by the individual that has spotted the nest, this can lead to issues with data transfer (from the application to the database) such as corrupted entries, missing values and problematic coordinates these were also identified and removed. Finally, there are issues with multiple replicated entries. Two different individuals could spot the same nest in two different days and that will reflect as two different data entries.

Both tables were checked for problematic entries using the OpenRefine software. The first data set consisted of 114 131 nests removed, corresponding to 113 952 nests with valid entries (i.e. non-zero and non-blank coordinates in the latitude or longitude field). The second data set include a total of 116 920 nests destroyed and this number is reduced to 116 508 valid entries. Subsequently both sets were merged into one eliminating replicated records.

The YLH sightings were loaded into QGIS and transformed into points. Next, to check for replicated entries at a spatial level, identified by different individuals, a buffer, 4m in radius, was created around each of the nests. This radius represents the minimum distance between two distinct YLH nests. This assumption was based on the previous report that the average distance between the nearest nest is about 700m however as time progresses this distance will decrease to only a few meters (Monceau & Thiéry, 2017). These buffers were created, dissolved and given a unique identifier. Next, the original sightings were intersected with the buffers and therefore each nest had a buffer id. The next step was standardising certain fields, the "observation type", "support", and a few other fields were edited so that the spelling for each category was consistent (Capital letters, plurals). Finally, the three fields (buffer id, "observation type" and "support") were concatenated into one. This was the field to be used to check for multiple replicated entries as two or more entries shared the same values in this field. That means two nests that are located less than 4m apart, on the same structure and exterminated on the same day can be treated as multiple replicated records.

From a total of 233 428 entries, 169 714 were found to be multiple replicated records of these only one per replicate was left and as such the final number of nest sightings for continental Portugal between the years of 2013 and 2023 are 143 758.

Once the data curation was finished, the YLH data was separated by year and imported into Quantum GIS (QGIS) where they were projected as points using the coordinates provided as per the flowchart in annex 1.

2.4. Landscape data

Once the study area was isolated, the next step was clipping both the land use data and the burnt area data into the study area to obtain the "COS_Study_Area" (flowchart annex 2) and the "BA_Study_Area" (flowchart annex 3) respectively. The same was done to the YLH nest sightings to obtain the sightings within the area of interest. A 1000m buffer was also applied to the original study area and clipped with the land use and the burnt area data. This was done to guarantee that the buffers generated in nests close the district boundaries would still be usable in the "Nest location analysis".

The burnt areas from 2019 to 2023 were selected, exported and clipped into the study area. A merged version of these years was created providing the total burnt areas for the study area for the time frame.

The land use charter consisted of a total of 83 classes that were aggregated into 8. The criteria that define each of the 8 classes can be found below:

> *"*Artificial areas*"* – Any area covered by anthropogenic structures regardless of use. Including factories, farmhouses and parking lots.
> "Transport routs" – Impermeabilized roads and highways with a width greater than 20m as well as any gas stations and rest stops (so long as these aren't included in urban fabric).
> "Shrubland" – Areas with a vegetation cover greater 10% and tree cover bellow 10%.
> "Permanent crops" – Actively managed olive groves, vineyards and orchards.
> "Hardwood forests" – Areas with a total hardwood tree cover above 10% as well as any hardwood agroforestry systems.
> "Conifer forests" – Areas with a total conifer tree cover above 10% as well as any coniferous agroforestry systems.
> "Temporary crops and pastures" – Fields covered by temporary crops at least once in the past five years, include rice fields and greenhouses as well as pastures used for grazing or fodder.
> "Areas without vegetation" – Areas with less than 10% vegetation cover including rocky outcrops, bodies of water and rivers.

2.5. Data Analysis

2.5.1. Yellow Legged Hornet nest density

For this part of the study, the interaction between the density of YLH nests and multiple variables was studied at a municipality scale. In total, 96 municipalities from 6 districts were used. Due to the inconsistent nature of the YLH sightings the variable density of YLH nests was calculated using the values for the year with the highest number of sightings for that municipality. The density was calculated in square kilometres for all variables.

The same land use classes used in the nest location analysis was also used for this analysis. Also included was the human population density for the year of 2019.

Other variables used were the density of honeybee colonies and apiaries and burnt area. These variables were collected at a municipality scale for the year with highest YLH nest sighting.

The land use data was clipped into each municipality and the area of each land use class was recalculated. This was exported into excel where the sum of each land use class per municipality could be determined and its respective percentage within the municipality was established.

2.5.2. Local nest position

For this analysis a buffer (250m radius) was generated around each nest location. The buffers were then intersected with the LU data as well as the data pertaining to the burnt area. The proportion of each of the 8 LU classes (flowchart annex 4) as well as burnt area (flowchart annex 5) was calculated for each buffer.

Out of all the nests in the study area, sighted between 2013 and 2023, 500 were chosen for each of the six districts. Thus, a total of 3000 buffers were randomly selected using the “Random selection” function in QGIS. These 3000 buffers constituted the subset of true nest sightings.

To obtain the pseudo-absence, a negative of the study area was generated with the buffers for the true sightings. A 500m internal buffer was then created to guarantee the pseudo-absence buffers did not overlap with the true sighting ones. Using the “random points inside polygon” tool a further 500 pseudo-absences were generated at random, at least 2000m apart. This was then repeated to all six districts in the study. In total, 3000 pseudo-absences were collected. The final sample size was 6000 buffers, half of which were true and the other half pseudo-absences.

2.5.3. Trends in Beehive

For this study, the main objective was to analyse the potential impact of the Yellow Legged Hornet has on the trend of honeybee colonies and apiaries. Since the invasion process was not homogenous throughout the country and some municipalities have felt the effect of the YLH for longer, two predictors were used; i) The average number of hornets observed from the year after the YLH was first detected in that municipality up to 2023, and ii) the maximum number of hornets nests in the same time span. The slopes used represent the variation of apiaries and colonies between 2018 and 2023. The year 2018 correspond to the year where the YLH was already completely distributed and well established in the region. Finally, four more linear models were done using burnt areas with the YLH and honeybee data in order to understand what the effect of this variable will be on both hornets and honeybees.

The intersection between the YLH sightings and the study area using geographic information systems (GIS) resulted in the maximum and average number of YLH sightings in each of the municipalities.

A total of four generalised linear models were applied on this data. Two of them analysed the effect the maximum number of YLH nests presence had on the variation of colonies and apiaries. The following pair of models used the average number of YLH sightings to study the variation on apiaries and colonies. Finally, the final eight models were used to understand the effects the burnt areas had on both colonies and apiaries as well as the maximum and average YLH values.

2.6. Models

For the first two studies (Hornet nest density at municipality level and Local nest position) a Boosted Regression Tree model (BRT) was used with a Gaussian distribution assumed for the first study and a Bernoulli distribution for the Local nest position. A Generalised Linear Model (GLM) assuming a normal distribution was used also for modelling the Hornet nest density at municipality level. In the YLH density analysis, the maximum density of YLH nests, corresponding to the year with the highest recorded number of sightings was used as the response. The predictors variables were: i) human population density (square kilometre) for 2019; ii) land use proportions for each category of LU within municipality-; iii) the proportion of burnt area for the year with most sightings per municipality, and iv) the density average of honeybee colonies and apiaries in square kilometres between the years of 2015 and 2023. For the local nest position analysis, we used the presence and pseudo-absence (randomly generated) of YLH sightings as the response variable while utilizing the same land use classes (within each buffer) as well as the proportion of burnt area as predictor variables.

For both BRT and GLM the first step in the modelling was checking the multicollinearity between the variables. Whenever two variables with high collinearity values above 0.7 were identified, then one was dropped (with less biological meaning).
Once the variables were dropped the initial model was fitted after which a simplification was attempted to determine if any of the remaining variables could be dropped without affecting the deviance of the model.

A linear model (LM) was used for the Trends in Beehives where eight different models were used to determine the impact caused by the YLH on local populations of farmed Honeybees. Four models used the average and maximum number of sightings as predictors and the decline in honeybee Colonies and Apiaries. While a further two models used wildfire occurrences for the region as predictors and the maximum and average number of sightings as response and finally two more models using the same predictor with honeybee Colonies and Apiaries as response. The predictors pertaining to YLH sightings was used to analyse the response decline of both the honeybee populations in the 5-year period between 2018 and 2023. While the wildfire predictors were used to understand if this impacted in any way the distribution of the YLH as well as the variation on honeybee populations.

3. Results

3.1. Influences on the density of nest sightings

With a clear distinction between the north and the centre region of the country it is important to understand how the sightings vary between these two regions. The number of nests observed agrees with the fact of a longer invasion period on the northern districts and shorter in the south following the expansion of the YLH. At a regional level the pattern is the same with a mean of 0.79 (±0.17) nests (Km^2) for the northern districts while in the centre the mean is of 0.18 (±0.06) nests (Km^2). Globally, for all six districts, the mean is 0.39 (±0.09) YLH nests (Km^2) (Fig 2).

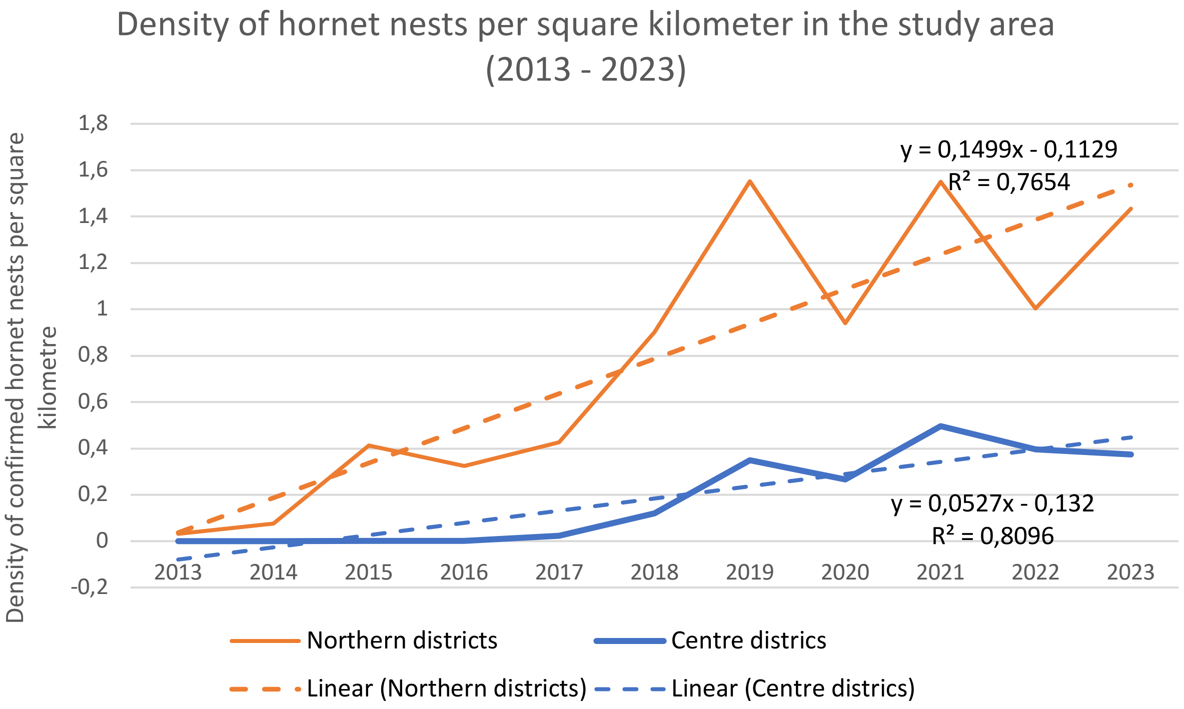


Figure 2 - Density of hornet nests (square kilometre) over the past decade (2013 – 2023) for the northern and centre districts of the study area

When looking at the study area, and the values used in this case the scenario changes slightly. The average number of nests for both the north and the south increases since the year with the highest number of sightings is used (table 1). The area with the highest proportion of nest sightings is in and around the Porto metropolitan area in the north with density ranging from 3 – 10 nests/km^2 (Fig. 3).

Likewise, when it comes to land use these two regions also vary quite a bit. While the main land use for each region is the same, “Hardwood forests” with a 4% difference in area between the regions, the differences begin to appear when looking at the other classes. The centre boasts around 12% more “Conifer forests” than the north making this the second most prominent land use class for the centre region. Meanwhile, in the north, the second most dominant land use are the “Temporary crops and pastures”. Finally, the last significant difference is the amount of “Artificial Area” varies a lot in each region, the north has around 14.5% of its area dedicated to this class while the centre has closer to 6%. For further information consult annex 6.

The north has a higher human population density at 399 people (Km^2) while the centre has a lesser density at around 92 (Km^2). The proportion of burnt area was slightly higher in the north with around 5.9% against the centre with 0.1%.

Finally, the density of honeybee colonies is lower in the centre with around 3.68 colonies/km^2 and 0.40 apiaries/km^2, meanwhile the north only has around 8.22 colonies/km^2 and 0.44 apiaries/km^2.

Table 1 - Distribution of the average (± standard error) density (N/km2) of hornet nest sightings per district

		North	
	Porto	Braga	Viana do Castelo
Hornet nest density	3.35 ±0.50	3.05 ±0.52	1.68 ±0.25
Municipalities	18	14	10
		Centre	
	Leiria	Santarém	Coimbra
Hornet nest density	0.87 ±0.11	0.43 ±0.12	1.21 ±0.13
Municipalities	16	21	17

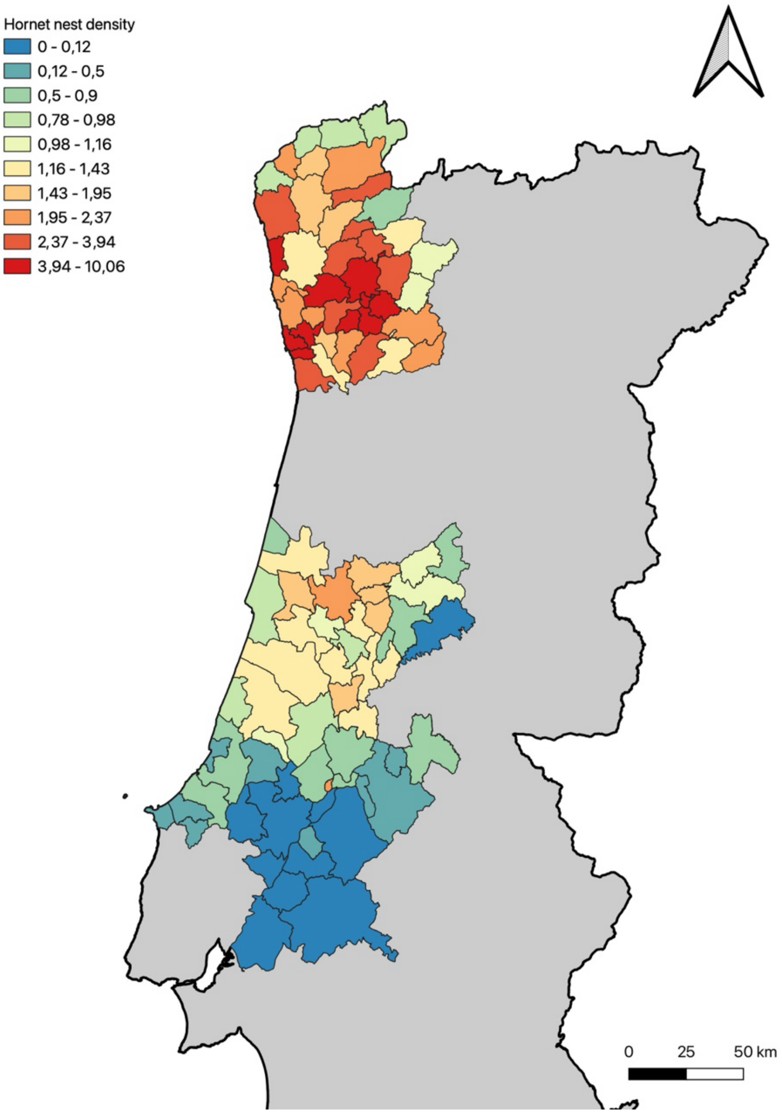


Figure 3 - Distribution of the density (N/km2) of hornet nest sightings per municipality using the year with the highest number of sightings per municipality between 2013 and 2023

3.1.1. Influences on the density of nest sightings – Results BRT Global

The multicollinearity matrix identified three pairs of predictors with values exceeding the 0.7 threshold: “apiaries” and “colonies” and “Artificial areas” with the “Population density” and “Transport routs” (Fig. 4). Since “Artificial areas” cover only about 9% of the study area, it was removed. Additionally, “colonies” were retained because they are more likely to vary annually than the apiaries.

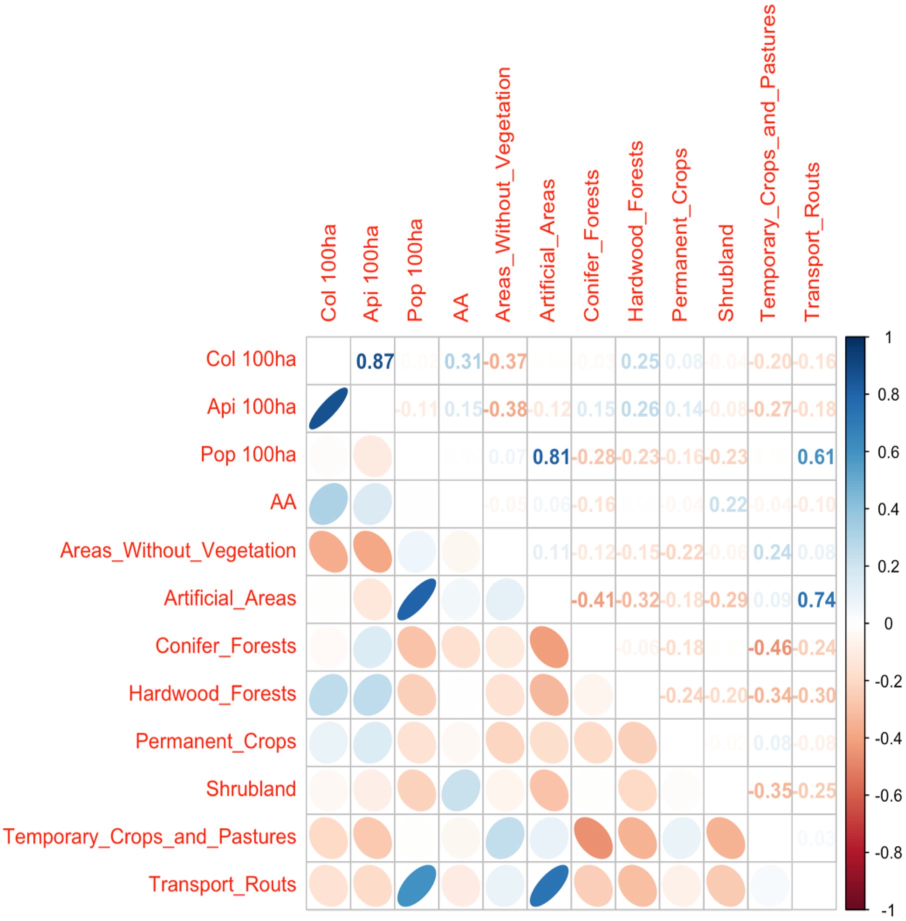


Figure 4 - Multicollinearity matrix for the variables on all 96 municipalities

After removing the aforementioned predictors, the model was run with 10 predictors, a tree complexity of 2, a learning rate of 0.007 and a bag fragmentation of 0.75. This resulted in 1650 trees with a cross-validation deviance of 1.59, a training correlation of 0.95 and a cross-validation correlation of 0.67 (±0.071).

Table 2 illustrates the importance of each variable. The “Population density” was the most influential predictor (45.3%) (Fig. 5 – a) followed by “Conifer forests” (14.7%) (Fig. 5 – b) and “Honeybee colonies” (10.2%) (Fig. 5 – c). Interaction analysis between the variables yielded “Hardwood forests” and “Population density” as having the largest interaction (3.82) followed by “Population density” with “Honeybee colonies” (2.31) (Fig. 6) and, finally, “Temporary crops and pastures” with “Honeybee colonies” (2.00). The number of YLH nests increased with population density and “Honeybee colonies”, whereas it decreases with the proportion of Conifer Forest. As for the temporary crop the pattern does not reveal a clear trend (Fig. 5 – d).

Table 2 - Relative influence of each predictor for the BRT of the 96 municipalities

Predictor	**Relative Influence**
Population density	45.33
Conifer forests	14.71
Honeybee Colonies	10.17
Temporary crops and pastures	6.74
Hardwood forests	5.89
Shrubland	4.99
Transport routs	3.89
Permanent crops	3.39
Burnt Areas	3.24
Areas without vegetation	1.65

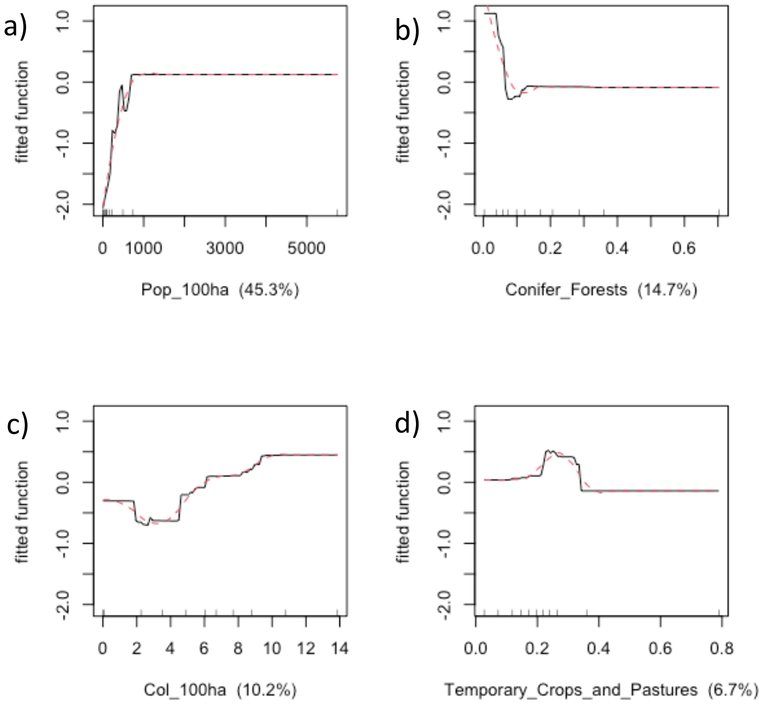


Figure 5 - Partial response graphs for the four most influential predictors for all 96 municipalities

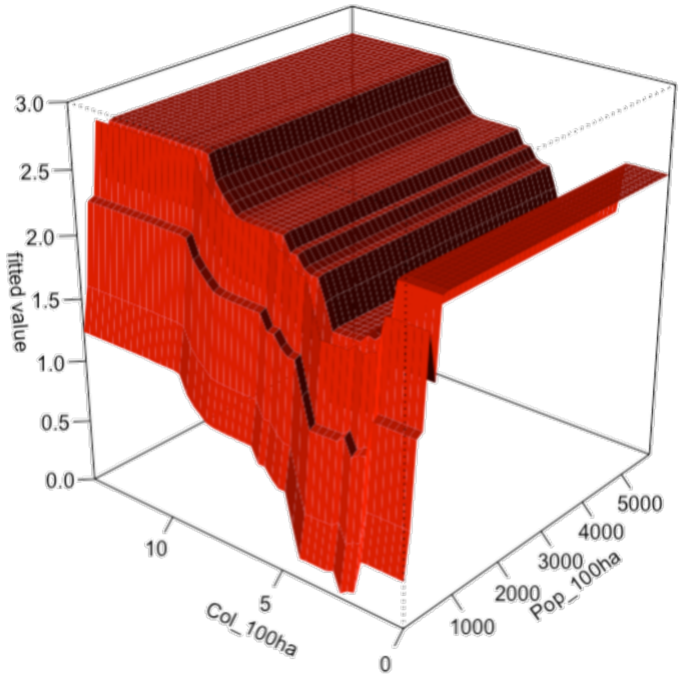


Figure 6 - Response of interaction between honeybee colony density and population density

1.1.1. Influences on the density of nest sightings – Results LM global

The first step was analysing the distribution of the variables and choosing those that had to be normalised (through a log operation) due to their skewed distribution. The following graphs (Fig. 7) illustrate the distribution of the 10 selected predictors.

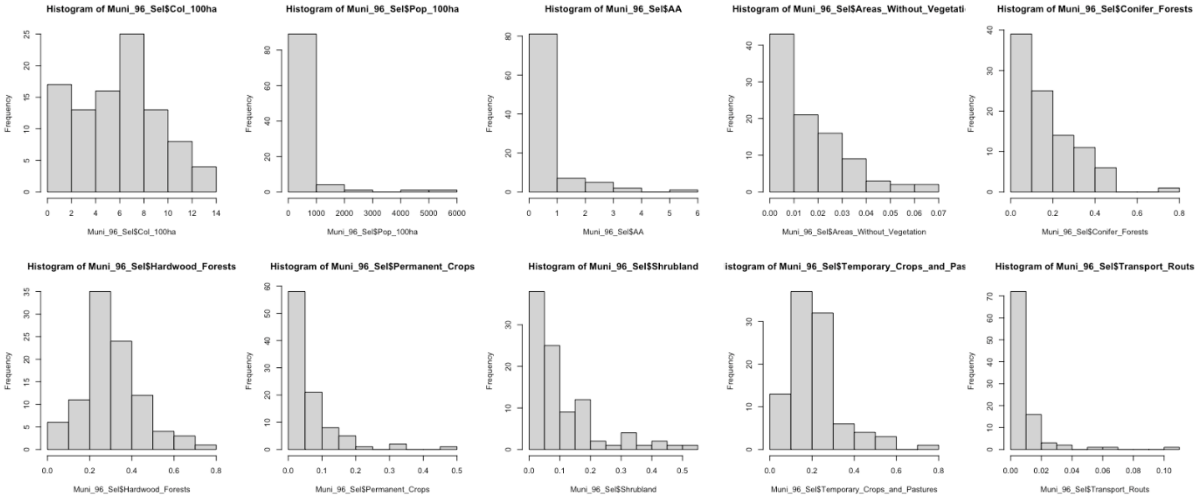


Figure 7 - Distribution of the selected predictors, pre-normalization, used for the linear model of the 96 municipalities

Based on this distribution, all except “Honeybee colonies” and “Hardwood forests” were normalised.

Due to their reduced sum of weights in the saturated model both “Burnt areas” and “Population density” were not used in the final model. The coefficients, standard error and P-value for each predictor is illustrated in table 3. While honeybee colonies had a significant positive effect on the YLH density all other variables exhibited a negative effect.

Table 3 - Coefficients, standard deviation, and p-value for each predictor of the linear model

	Coefficient	**P-value**
Honeybee Colonies	0.12 (±0.03)	0.000578
Hardwood forests	-10.87 (±1.49)	1.36e-10
Areas without vegetation	-75.58 (±19.77)	0.000247
Conifer forests	-33.32 (±4.13)	3.83e-12
Permanent crops	-35.49 (±4.80)	8.43e-11
Shrubland	-32.54 (±4.61)	3.83e-10
Temporary crops and pastures	-30.51 (±5.22)	9.04e-08
Transport routs	-60.06 (±30.61)	0.052956

Out of all the eight predictors, “Honeybee Colonies” is the only variable that has a positive, influence in the detection of the YLH (Fig.8). “Areas without vegetation” is the variable with the strongest effect (in this case negative). Finally, “Transport routs”, despite having a is the only one with no statistical significance.

Both the Multiple R-Squared (0.6996) and the Adjusted R-Squared (0.6393) indicate a good fit for the model. The overall model also appears to show statistical significance (F-Statistic = 22.04, P-Value < 2.2e-16).

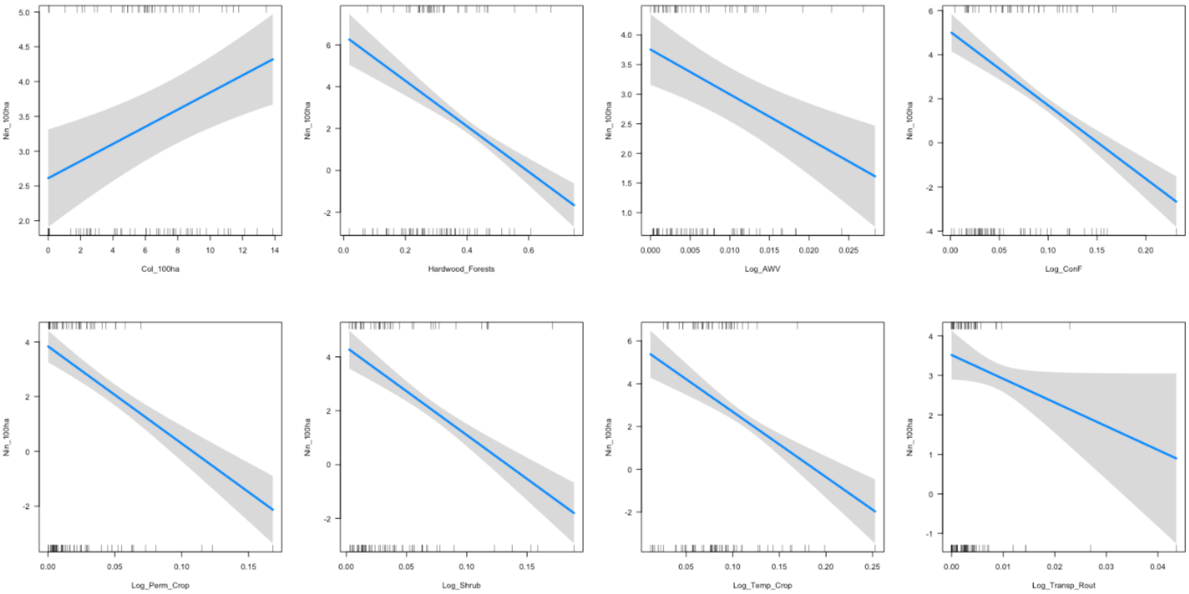


Figure 8 - Response graph for each of the predictors for the linear models of the 96 municipalities

1.1.2. Influences on the density of nest sightings – Results BRT Centre

The first multicollinearity matrix (Fig. 9) yielded similar results as the one done for all 96 municipalities. The only difference is the pair “Population density” with “Transport routs”. As the latter seemed to have little influence when analysing the totality of the study area this variable was also dropped.

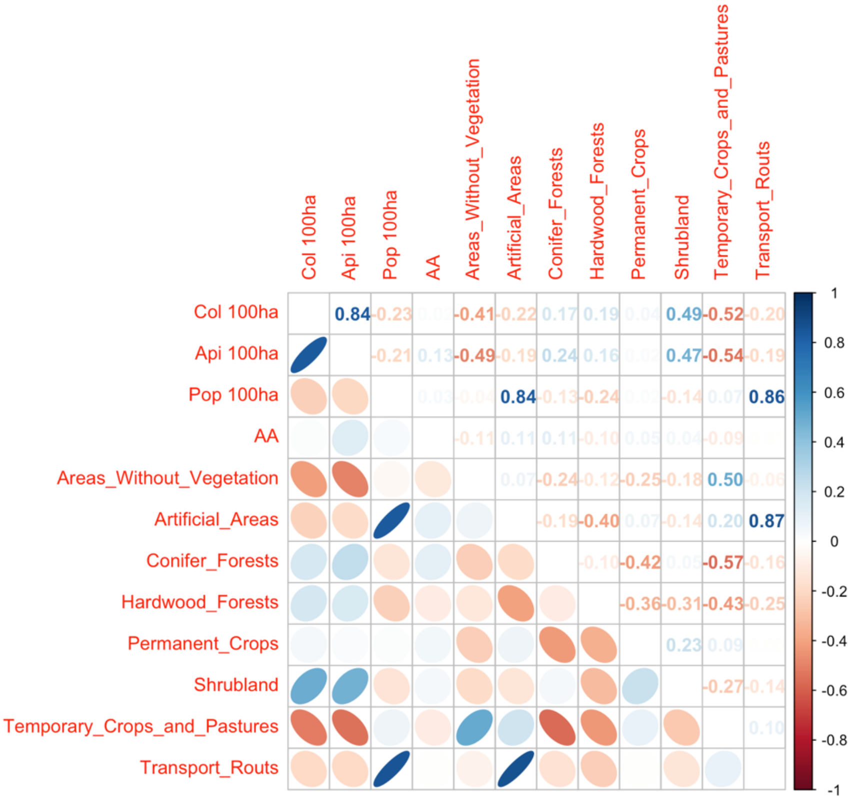


Figure 9 - Multicollinearity matrix for the predictors used in the BRT for the 54 municipalities in the centre

After removing the predictors needed the model was run with 9 predictors, a tree complexity of 2, a learning rate of 0.002 and a bag fragmentation of 0.75. This resulted in 1100 trees with a cross-validation deviance of 0.299, a training correlation of 0.809 and a cross-validation correlation of 0.43 (±0.164).

When looking at the relative influence for each predictor, the centre varies a fair bit from the entirety of the 96 municipalities. As illustrated by table 4 the main predictor in this case is “Conifer forests” (38.4%) (Fig. 10 – a) followed by “Shrubland” (12.7%) (Fig. 10 – b), “Burnt areas” (12.5%) (Fig. 10 – c) and finally “Areas without vegetation” (11.6%) (Fig. 10 – d).

No major significant interaction was observed as they all had interaction sizes less than 1.

Table 4 - Relative influence of each predictor for the BRT of the 54 municipalities of the centre region

Variable	**Relative Influence**
Conifer forests	38.39
Shrubland	12.73
Burnt areas	12.49
Areas without vegetation	11.61
Honeybee colonies	10.67
Population density	8.28
Hardwood forests	2.42
Permanent crops	2.23
Temporary crops and pastures	1.18

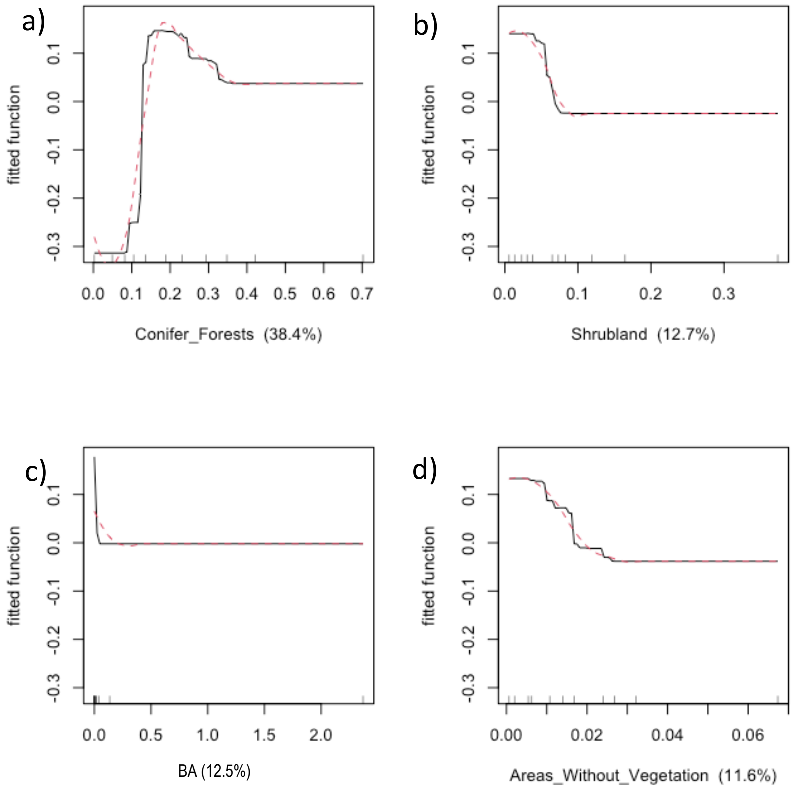


Figure 10 - Partial response graphs for the four most influential predictors for all 54 municipalities for the centre

1.1.3. Influences on the density of nest sightings – Results LM Centre

Similarly to the global results, all but two variables do not need to be normalised (Fig. 11), “Honeybee colonies” and “Hardwood forests”. The remainder were normalised.

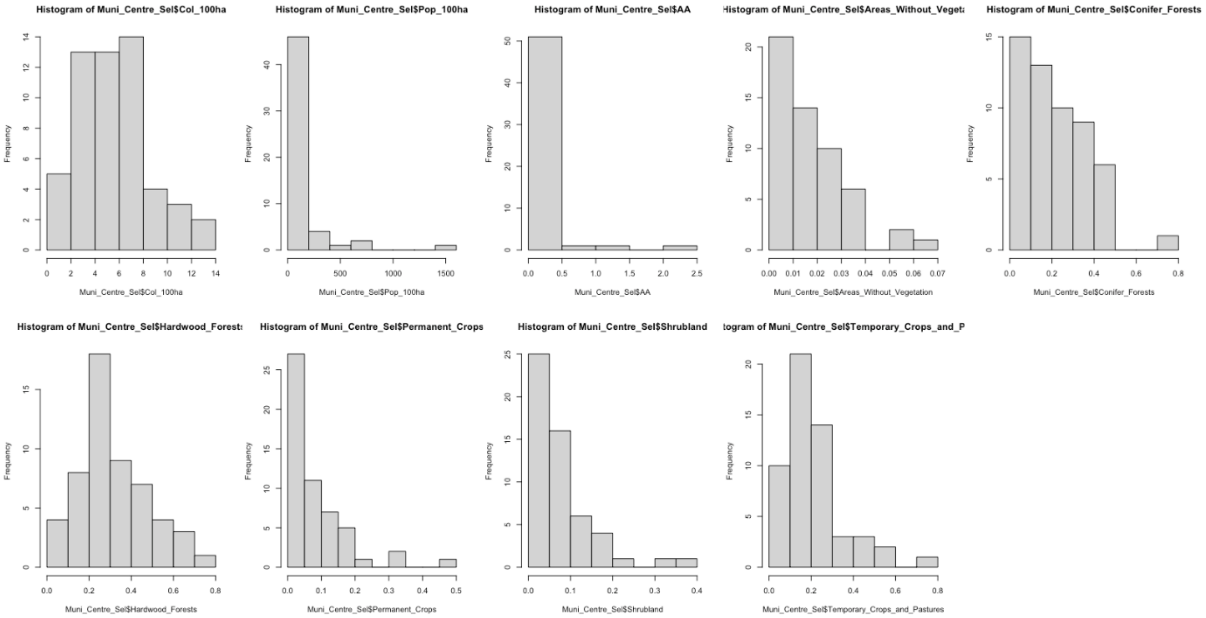


Figure 11 - Distribution of the selected predictors, pre-normalization, used for the linear model of the 54 municipalities of the center

The final model utilized three predictors “Areas without vegetation” (AWV), “Permanent crops” and “Population density” (Fig. 12). Of these, “Areas without vegetation” appear to have a negative influence with a coefficient of -44.89 (±11.98) with a p-value < 0.001. The “Permanent crops” also seem to have a negative influence (albeit not as much as AWV) with a coefficient of -7.1 (±2.2) and a p-value < 0.01. Finally, the “Population density” is the only predictor with a positive coefficient at 0.43 (±0.2) and a p-value < 0.01. The final model had an R-squared 0.32, the F-statistic is 7.81 with a p-value < 0.05.

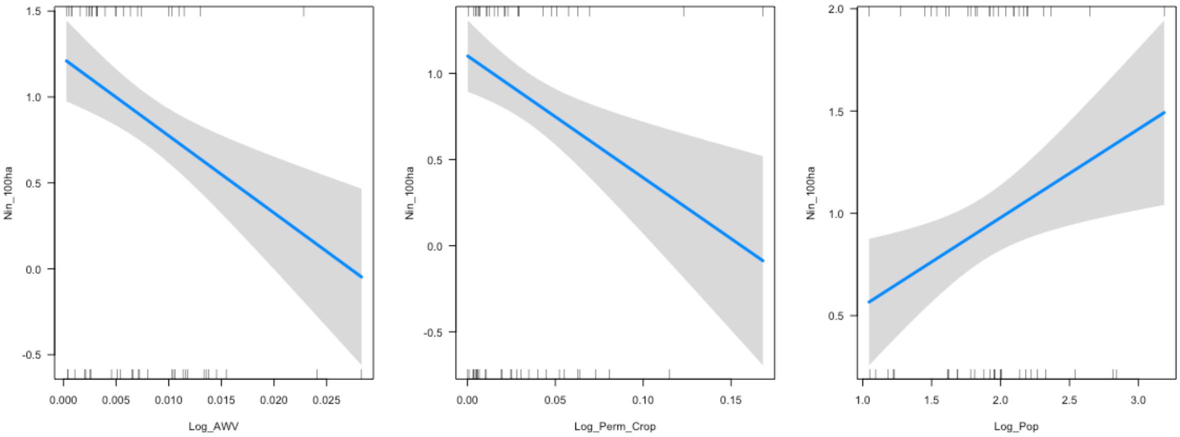


Figure 12 - Response graph for each of the three predictors used in the linear models of the 54 municipalities

1.1.4. Influences on the density of nest sightings – BRT Results North

The multicollinearity matrix for the North is very similar to the one for the centre. The same pairs of variables yielded multicollinearity values over 0.7 albeit with slightly different values (Fig.13). That being the case, the variables selected are the same as the ones in the Centre.

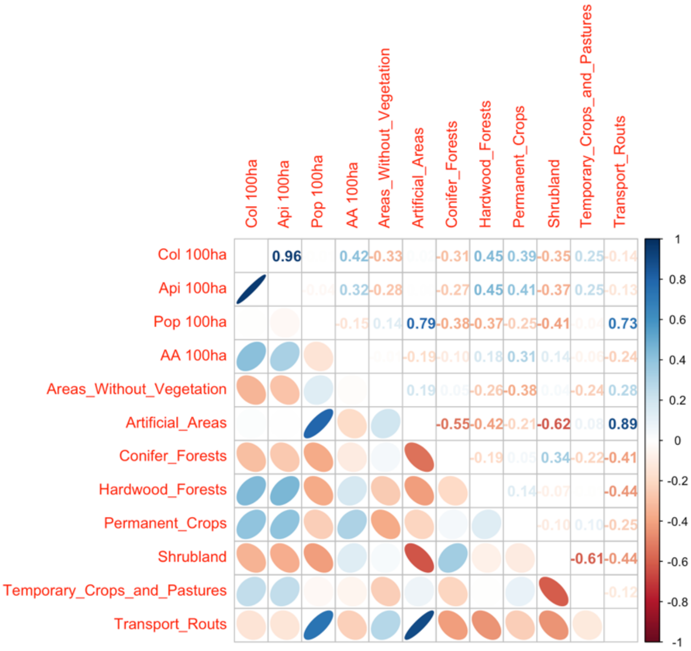


Figure 13 - Multicollinearity matrix for the predictors used in the BRT for the 42 municipalities in the north

After the selection of the predictors the model was run with a tree complexity of 2, a learning rate of 0.003 and a bag fragmentation of 0.75. This resulted in a model with 1200 trees with a cross-validation deviance of 2.654, a training correlation of 0.874 and a cross-validation corelation of 0.625 (±0.134).

When it comes to the relative influence of the predictors, the first two are the same for both regions, “Conifer forests” (33.2%) (Fig. 14 – a) and “Shrubland” (22.2%) (Fig. 14 – b). However, in this case, “Population density” was next with a weight of 22% (Fig. 14 – c). None of the remaining predictors had relative influences greater than 6% (table 5).

Finally, the interactions observed between variables were of negligible size relative to the first model. “Hardwood forests” and “Population density” had the highest interaction size at 0.55.

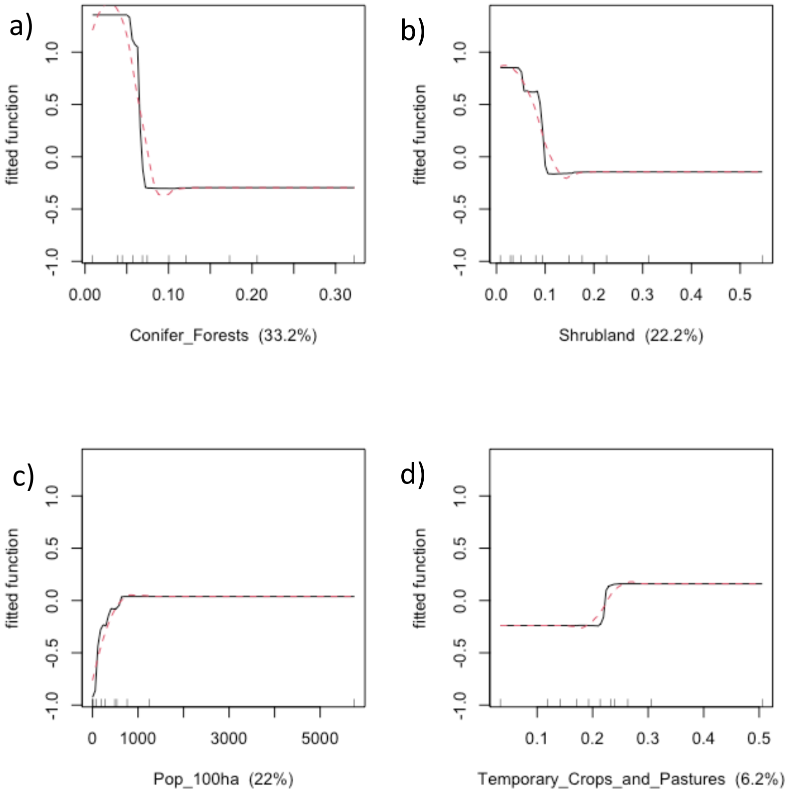


Figure 14 - Partial response graphs for the four most influential predictors for all 42 municipalities for

Table 5 - Relative influence of each predictor for the BRT of the 42 municipalities of the northern region

Predictors	**Relative Influence**
Conifer forests	33.16
Shrubland	22.20
Population density	22.00
Temporary crops and pastures	6.17
Hardwood forests	6.00
Honeybee colonies	4.74
Burnt Areas	3.18
Permanent crops	2.09
Areas without vegetation	0.46

1.1.5. Influences on the density of nest sightings – Results LM North

Just as it was done for the other linear models, some variables had to be normalized. In this case, all but two predictors were normalized, “Honeybee colonies” and “Temporary crops and pastures” as can be seen in figure 15.

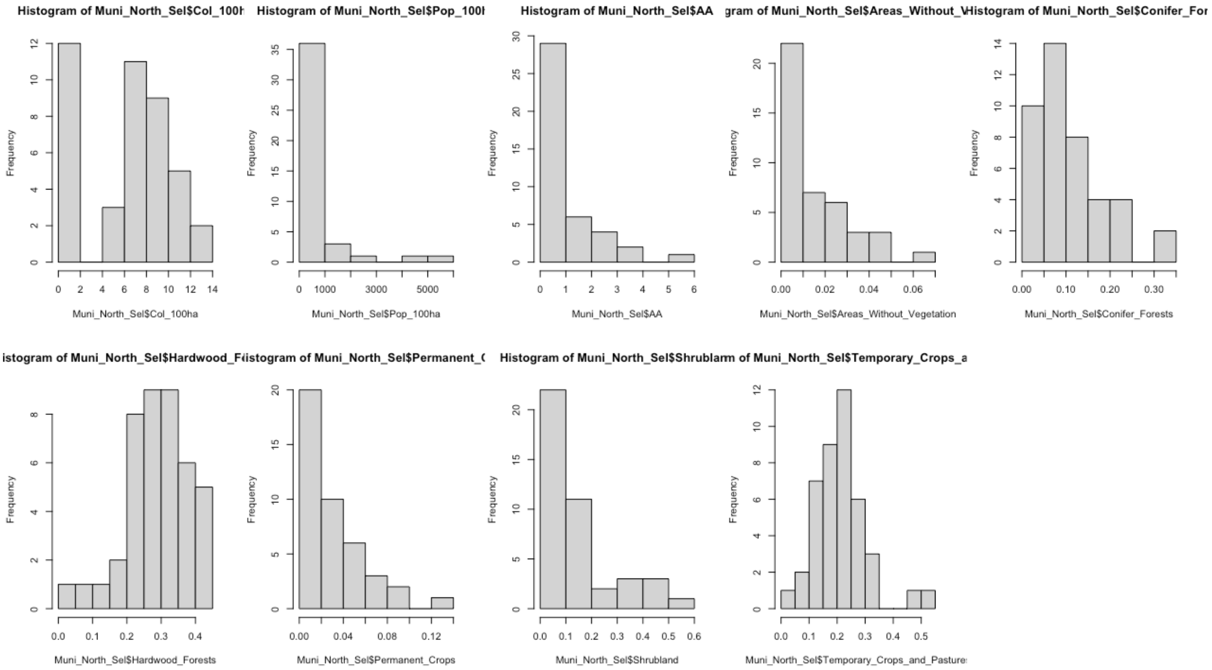


Figure 15 - Distribution of the 9 predictors used for the LM of the 42 northern municipalities

The final model was made using 4 predictors, “Honeybee colonies”, “Conifer forests”, “Hardwood forests” and “Shrubland”. Of these, “Honeybee colonies” presented a positive influence with a coefficient of 0.18 (±0.06) and a p-value < 0.01 (Fig. 16 – a). “Conifer forests” had a coefficient of -35.3(±7.8) and a p-value < 0.001 (Fig. 16 – b). “Hardwood forests” was the second most significant influence in this region with a coefficient of -33.83 (±7.12) and a p-value < 0.001 (Fig. 16 – c). Finally, “Shrubland” had a coefficient of -14.25 (±4.71) and a p-value < 0.01 (Fig. 16 – d). Overall, the model appears to have a reasonable fit with an R-squared of 0.69, an F-statistic of 20.95 and a p-value <0.001.

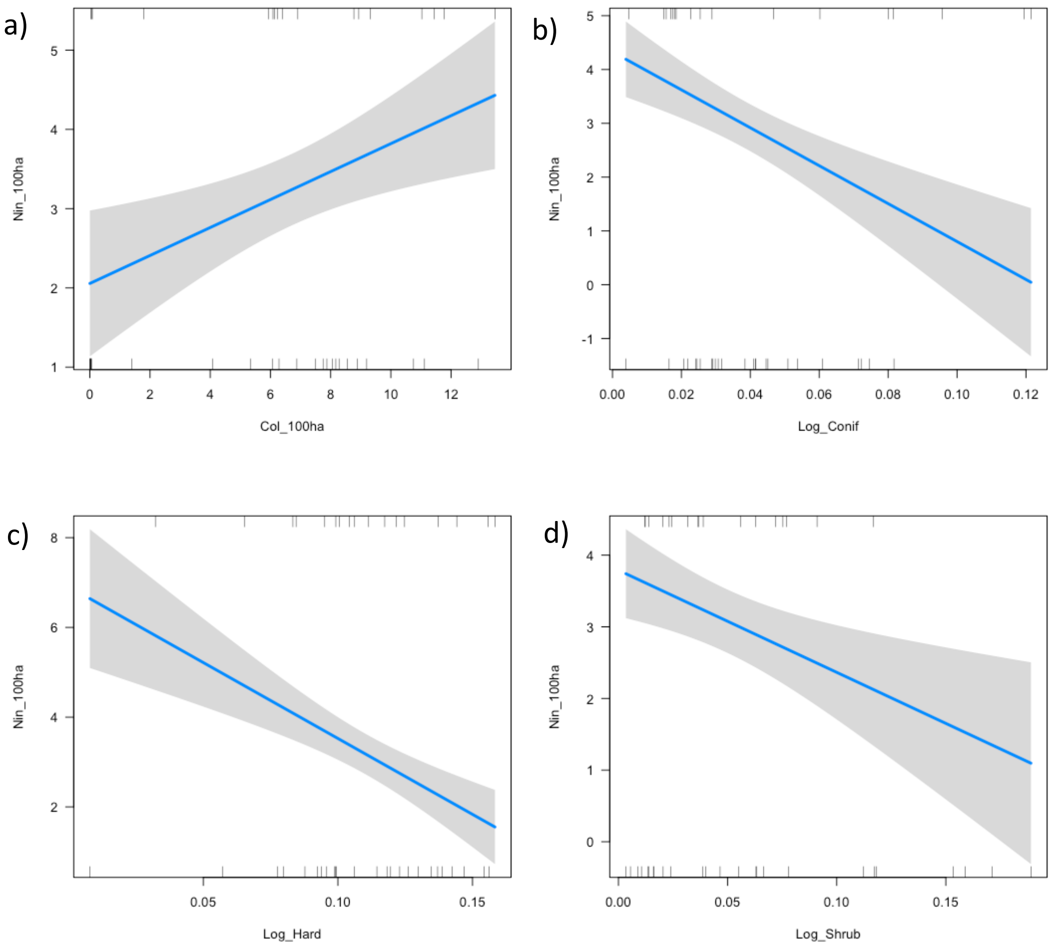


Figure 16 - Partial response graphs for each of the predictors used in the linear model for the 42 municipalities in the north

1.2. Nest location vicinity

1.2.1. Nest location vicinity – Results BRT

The multicollinearity matrix (Fig. 17) for the predictors did not identify any pair of variables with values surpassing the 0.7 threshold. As such, the analysis proceeded with all 9 variables.

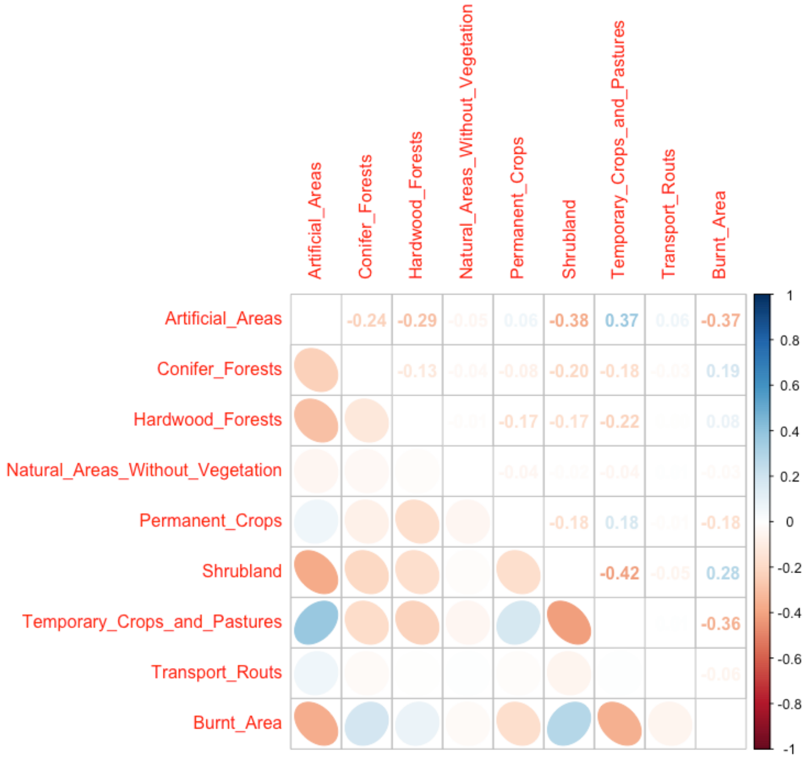


Figure 17 - Multicollinearity matrix for the predictors used in the BRT for the Nest location vicinity

A boosted regression tree model was then constructed with a binary response of true or pseudo-absence. The model had a tree complexity of 2, a learning rate of 0.013 and a bag fragmentation of 0.75. The model stabilized at 1650 trees with a cross-validation deviance of 0.609 (±0.014). The model achieved a training AUC score of 0.954 and a CV AUC score of 0.943. The model seems to have a good fit with the training data with a residual deviance of 0.554 and a correlation of 0.817.

A simplification was attempted to remove variables of lesser influence however, no predictors were dropped before the predictive deviance rose. As such, no simplification was used.

An analysis on the importance of each variable revealed a strong influence from “Artificial areas”, contributing around 82% (Fig. 18 – a). The next predictor with a strong influence were “Temporary crops and pastures” (9.1%) (Fig. 18 – b) followed by “Shrubland” (2,7%) (Fig. 18 – c) and “Hardwood forests” (2.1%) (Fig. 18 – d). Table 6 displays the remainder predictors and their relative influence. The partial response curve for the five most influential predictors reveals that all but “Shrubland” have a positive influence.

Table 6 - Relative influence of each predictor for the BRT of the nest location vicinity

Variable	Relative Influence
Artificial areas	81.98
Temporary crops and pastures	9.13
Shrubland	2.73
Hardwood forests	2.18
Conifer forests	1.98
Permanent crops	1.01
Areas without vegetation	0.42
Burnt areas	0.30
Transport routs	0.26

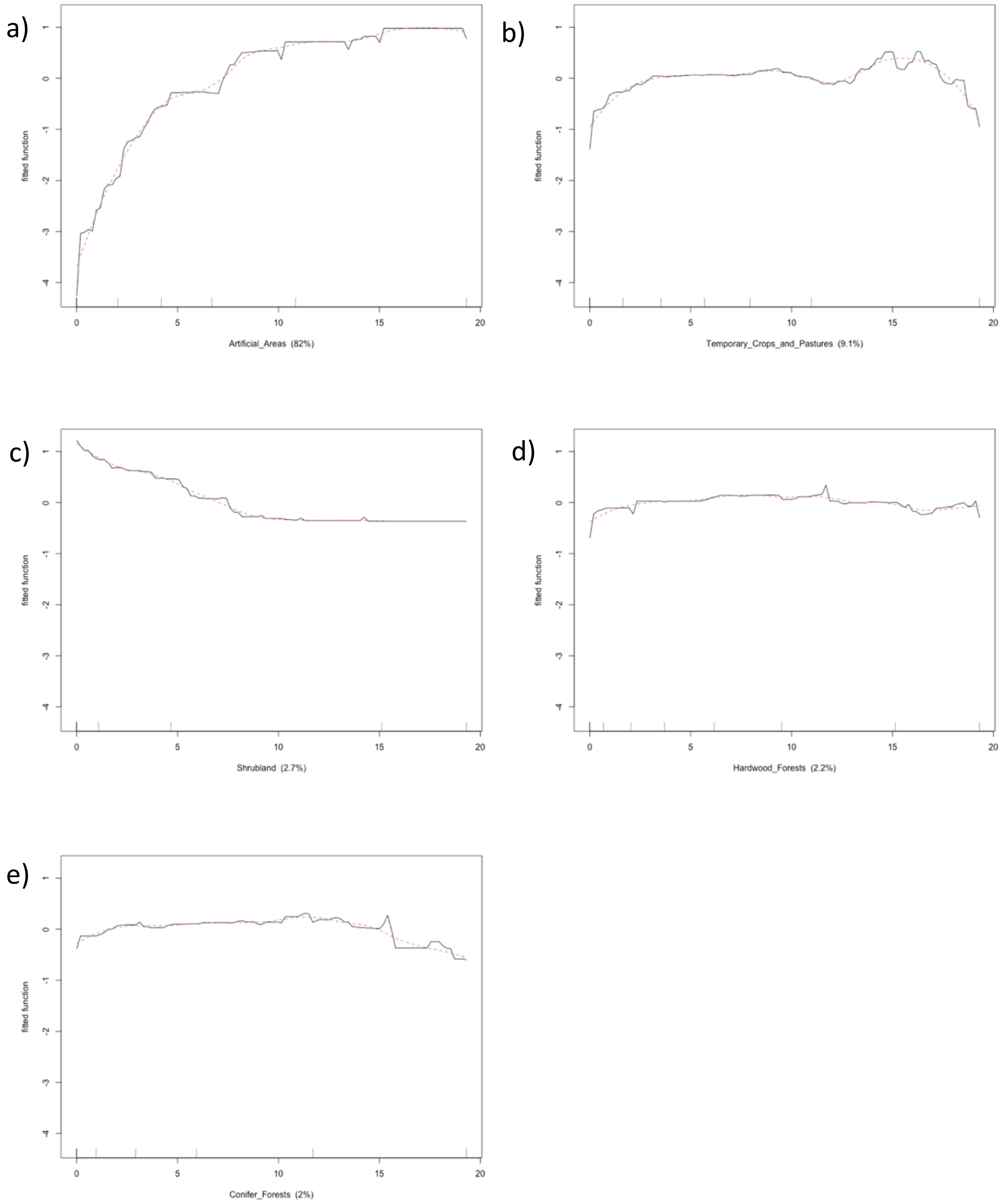


Figure 18 - Partial response for the four most influential predictors in the nest location vicinity BRT

The interaction between predictors (table 7) was also studied and the results are presented below. “Permanent crops” and “Conifer forests” being the interaction to most influence the model with an interaction size of 13.08. The next pair of predictors is “Transport routs” with “Artificial areas” (9.8) followed by the interaction between “Temporary crops and pastures” and “Hardwood forests” (9.11) and “Conifer forests” (8.45).

Table 7 - Interaction between predictors used in the nest location vicinity

Land Cover 1	Land Cover 2	Interaction size
Permanent crops	**Conifer forests**	13.08
Transport routs	**Artificial areas**	9.80
Temporary crops and pastures	**Hardwood forests**	9.11
Temporary crops and pastures	**Conifer forests**	8.45

1.2.2. Nest location vicinity – Results GLM

A Generalized Linear Model based on a binomial distribution and a logit function. was also applied to the buffer data. The distribution of the data was far from ideal, even after the logarithmisation of the data (Fig. 19). Due to high sum of weights (the lowest being "Burnt areas" with a sum of weights of 0.77) from the predictors and the fact that non had multicollinearity values above 0.7 all were used in the final model.

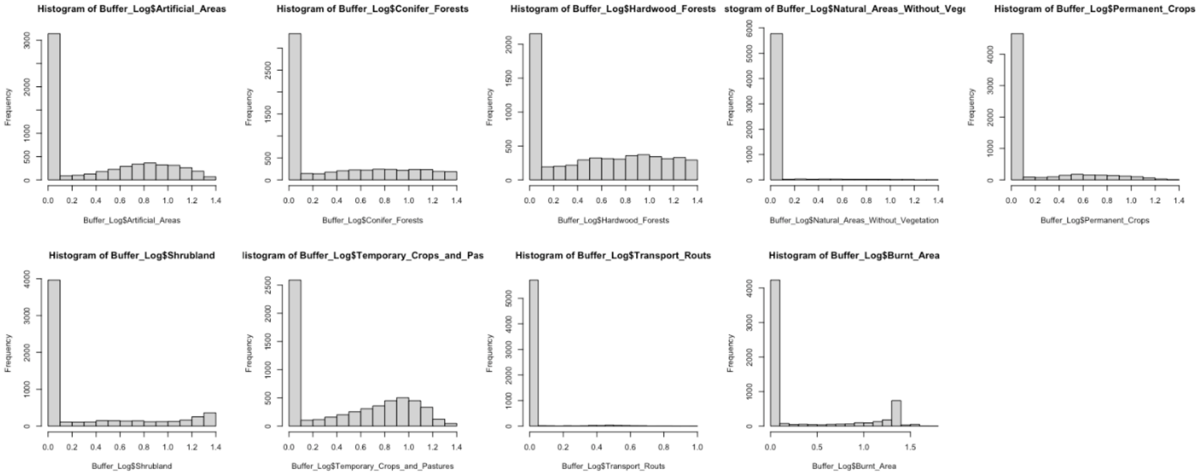


Figure 19 - Distribution of the values for each predictor used in the linear model for the nest location vicinity

Out of all the predictors used, "Artificial areas" is the main predictor ($\beta = 5.31$, p-value < 0.005) (Fig. 20 – a) and statistical significance followed by "Temporary crops and pastures" ($\beta = 2.51$, p-value <0.05) (Fig. 20 – d). Both "Conifer" (Fig. 20 – g) and "Hardwood" (Fig. 20 – h) forests seem to have similar positive influences in the outcome (β = 1.55 and 1.83 respectively, p-value < 0.05). Out of every predictor, "Shrubland" is the only one to have a negative relationship with the outcome ($\beta = -0.49$, p-value < 0.05) (Fig. 20 – c) corroborating the results. The adjusted R-square value of 0.72 suggests a good fit for the model.

These results seem to follow the same patterns as the BRT. Both point to buffers with high proportions of "Artificial areas", "Temporary crops and pastures" and both "Conifer" and "Hardwood" forests. The negative influence of "Shrublands" translated in both models with similar patterns.

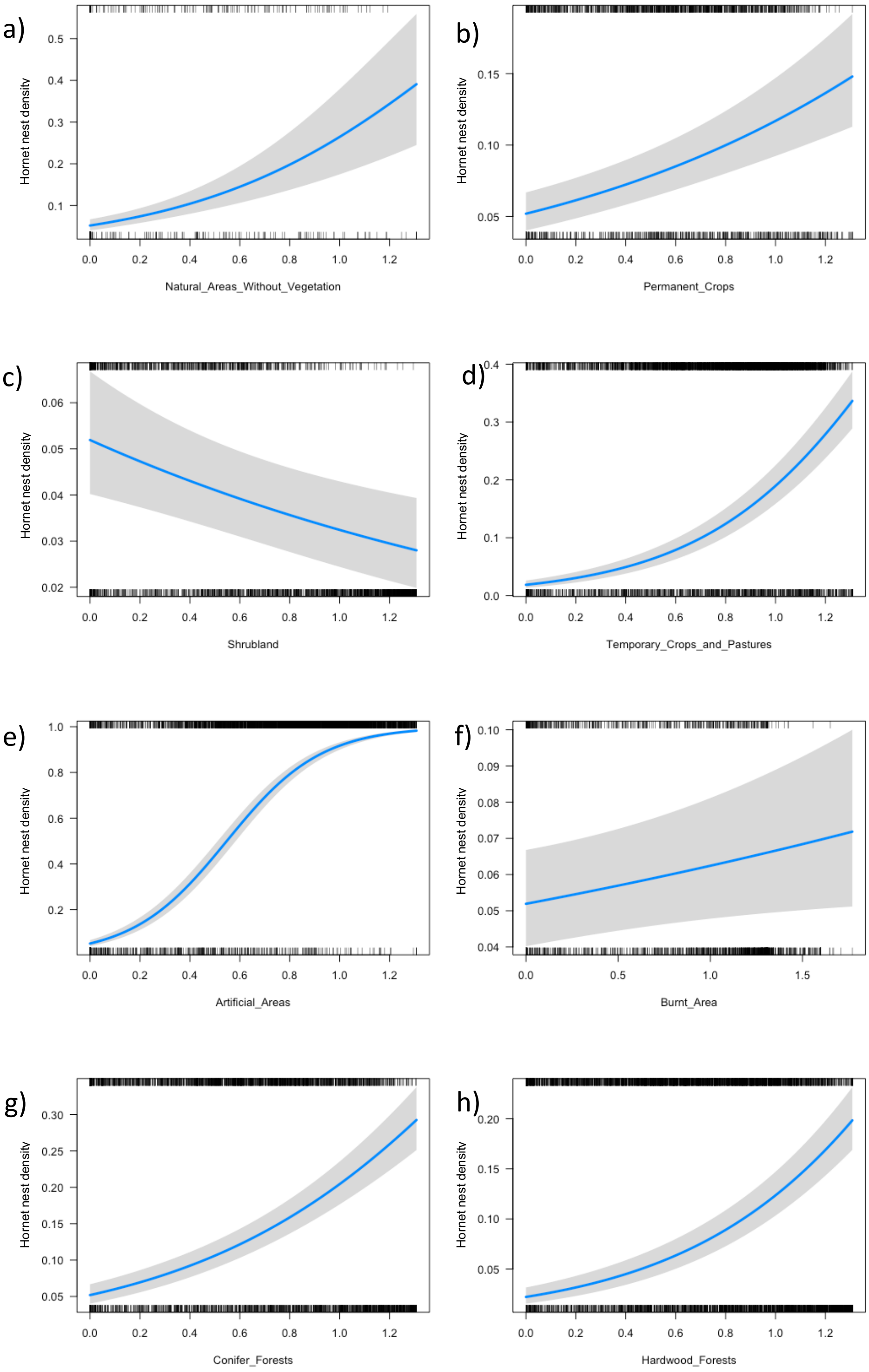


Figure 20 - Partial response for each predictor in the linear model

1.3. Trends in Beehive

The density of honeybee colonies does not appear to vary much between districts, on average the districts have around 8.22 honeybee colonies per square kilometer and around 0.44 apiaries per square kilometer. At a district level, similar values appear with the density for Apiaries in Porto (0.49 Apiaries/km^2), Braga (0.50 Apiaries/km^2) and Viana do Castelo (0.40 Apiaries/km^2) as well as the density of Colonies in Porto (8.84 Colonies/km^2), Braga (9.13 Colonies/km^2) and Viana do Castelo (8.37 Colonies/km^2).

The average slope on beehives (between 2018 and 2023) for the three districts indicates an overall decrease in the number of honeybee colonies) being significantly below 0 (T-test= -2.385, n=41, p=0.011) Braga is the one with the overall greatest decrease in the number of honeybee colonies, followed by Viana do Castelo and then Porto (table 8). When it comes to Apiaries, the average slope between the three districts is almost null, 0.87±13.21 and does not reveal a significant trend (T-test= 1.410, n=41, p=0.083). The highest slope is observed for Braga (table 8).

Table 8 - Average slopes (±standard error) for colonies and apiaries through the three study areas

Region	Colonies Mean	Apiaries Mean
Viana do Castelo	-38.36 ±(32.80)	0.16 ±(1.07)
Braga	-53.38 ±(23.98)	2.42 ±(0.96)
Porto	-10.73 ± (16.61)	0.08 ±(0.90)
Total, **n=41**	-32.77 ±(13.21)	0.87 ±(0.59)
t-test p-value	-2.385, p=0.011	1.410, p=0.083

The burnt area during the period of 2018 to 2023 varied between 5.26 and 6.81 % for the three districts (table 9).

Table 9 - Burnt area total and proportion

	Area district (km^2)	**Burnt area (km^2)**	**Proportion**
Viana do Castelo	2218,84	103,73	6.81
Braga	2706,15	89,74	5.69
Porto	2331,65	72,23	5.26
Total	7256,64	427,91	5.90

The decline in Apiaries in the northern region of the country was not significantly related with the maximum number of hornets observed (β = -0.0011, p-value = 0.564, R-squared = 0.008). Likewise, the variation in Apiaries was not significantly related to the average number of YLH nests (β = -0.0078, p-value = 0.120, R-squared = 0.054). The proportion of burnt area in a municipality also appears to have little influence in the number of Apiaries (β = 0.0753, p-value = 0.552, R-squared = 0.009).

However, the decline in the number of honeybee colonies over the analysed timespan can be explained, in part, by the maximum number of hornets observed (β = -0.123, p = 0.0049, R² = 0.181). When using the average number of nests (β = -0.218, p = 0.059, R² = 0.086) the significance is reduced, becoming borderline (Fig 2). Finally, the proportion of burnt area showed no significant relationship with the decline in honeybee colonies (β = 4.71, p = 0.09, R² = 0.0704) (Fig 21).

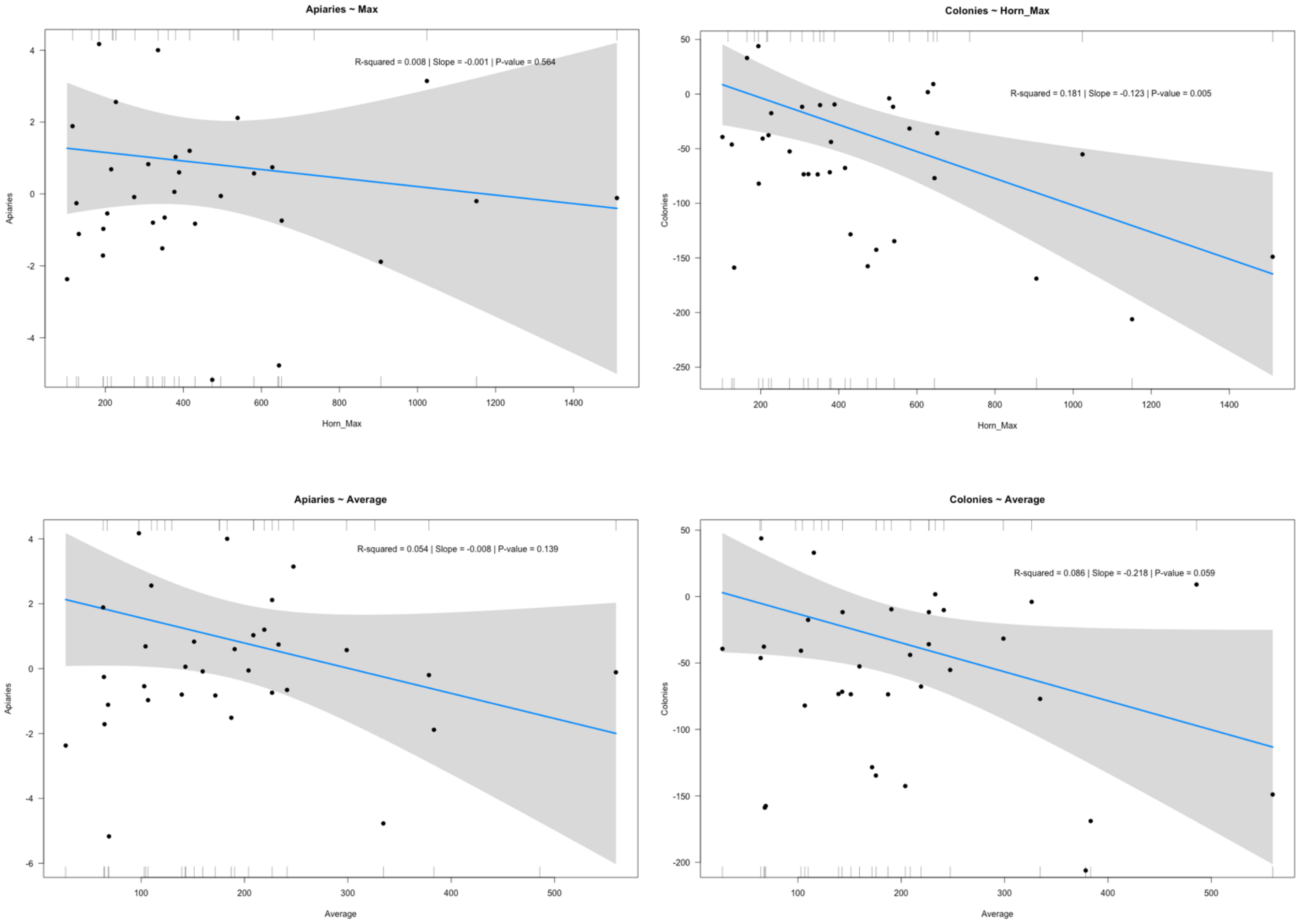


Figure 21 - Regression graphs for the linear models between hornet and honeybee populations

Two additional models were run to test the interaction between the maximum number of YLH nests (-6.97, 0.475, R-squared = 0.0129) and average number of YLH nests (-4.46, p-value = 0.236, R-squared = 0.035) with the burnt area. Neither model showed any statistical significance (Fig. 22).

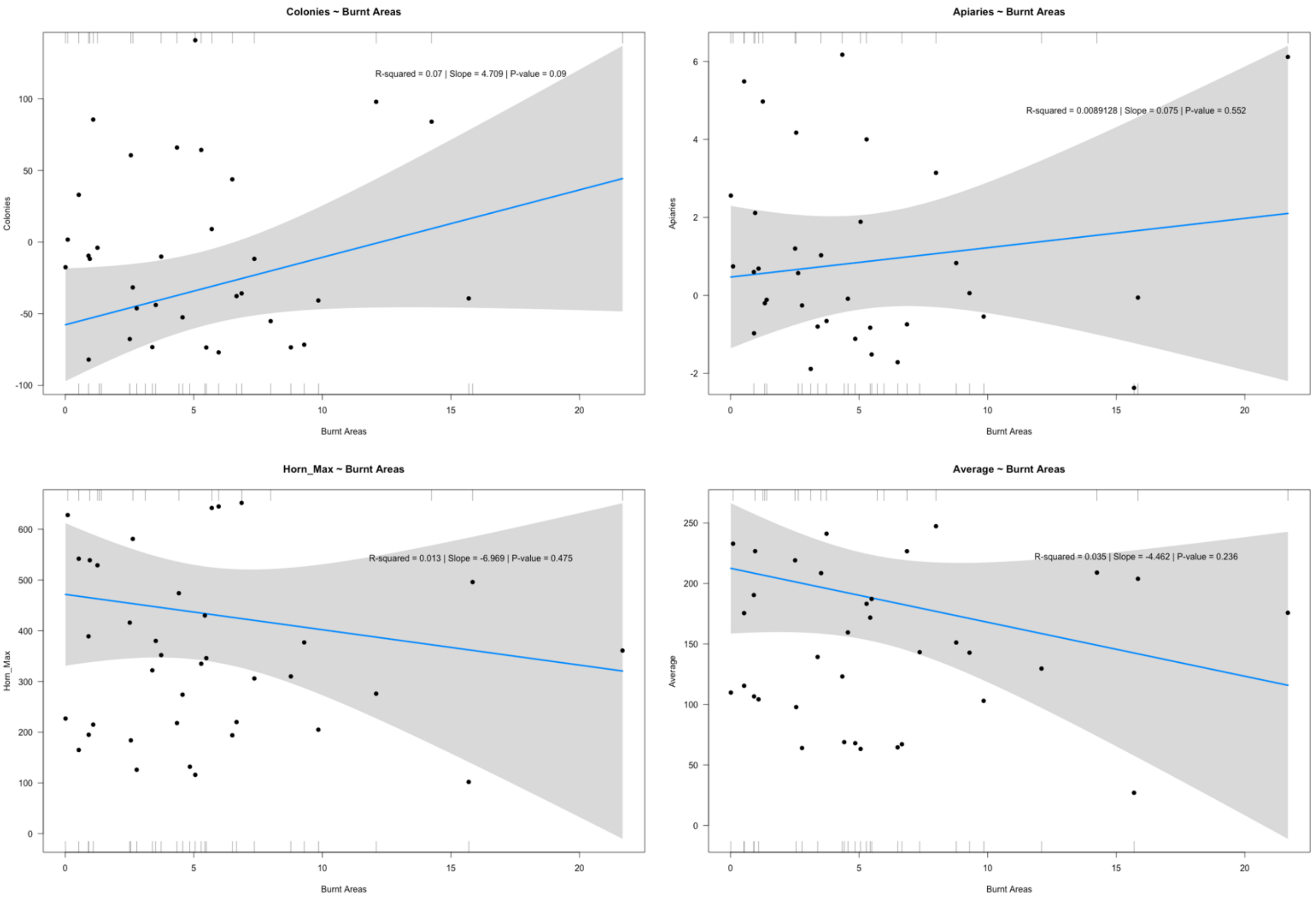


Figure 22 - Regression graphs for the linear models between the burnt areas and hornet/honeybee populations

2. Discussion

2.1. Influences on the density of sightings

Using the eight land use classes, human population density, wildfire occurrences and honeybee population values were used to understand how they influence the density of YLH observations.

The main factor that will influence the sightings of the YLH is the population density. This is in accordance with Johnston et al. (2020) statements about the biases that occur when using crowdsourced data.

However, “Conifer forests” are the land use class with the most significant influence, albeit negative, for the density of sightings despite having a smaller area proportion to “Hardwood forests”. This evidence is corroborated the findings by Catarino (2022) where the author concludes that this negative relationship may be caused, in part, by the poor biodiversity found in *Pinus pinaster* stands in Portugal. The same negative effect is noted on Hardwood forests however, not as pronounced as the conifers’ effect. This might happen because *Eucalyptus* stands represents the dominant hardwood in this type of landscape and tend to have also low biodiversity (Otero et al, 2015).

This remains the case even when analysing the centre and northern regions separately. Even in the case of the northern districts where “Hardwood forests” outweigh conifers by a considerable amount (annex 6). One possible explanation for this phenomenon is the fact that the main species of “Hardwood forests” in these regions is the *Eucalyptus*, specifically *Eucalyptus globulus* used to produce paper in monocultures with low biodiversity in the understory. Meanwhile, *Pinus pinaster* is the coniferous species most commonly found in Portugal (IFN 2015, https://www.icnf.pt/api/file/doc/c8cc40b3b7ec8541) capable of reaching substantial hights and having crown coverages around 10m in diameter. However, it is important to point out that the linear models point to a decrease in nest sightings the bigger the forested area is, regardless of type. This is another case of geographic biased caused by the nature of the data used, people are less likely to sight nests in large, forested areas because they are less likely to venture deep into these areas.

The presence of honeybee colonies was also selected as a driver to the number of sightings. Understandable as the YLH is known for being a prolific hunter of these animals. However, another reason for sightings to increase with the increase of beekeeping activities is the fact that the beekeepers tend to be particularly on the lookout for the YLH, as this may cause irreparable damage to the honeybee populations. Both linear models identified this as being one of the few factors with a positive impact on the sighting of the YLH (the other one being population density).

“Temporary crops and pastures” (frequently managed agricultural land) was the fourth most significant predictor when it comes to YLH nest sighting density affecting it negatively. This falls in line once again with the conclusions reached by Otero et al (2015). However, it contrasts with the conclusions reached by Bessa et al (2016), being that these types of areas have a positive influence on YLH nesting. These discrepancies might be in part due to the type of agriculture practiced by the locals. Large, extensive agricultural areas don’t have the appropriate structures for YLH nesting as well as diminished biodiversity levels (Norris, 2008). Still, in the North and Centre of Portugal agricultural parcels are of small size. Meanwhile, when looking at the fourth graph on figure 5, there is a small increase in nest densities until a certain threshold of the variable “Temporary crops and pastures”. This may be in part due to smaller parcels of agricultural land being close to nesting sites for the YLH as well as being local sources of biodiversity in areas surrounded by pure stands of forests.

Burnt areas appear to have little to no influence on the sightings of the hornets. The model used in the centre region of the country identified this as being the third most influential factor with a negative effect for the sighting of the YLH. However this may be in part because of the wildfires of 2017 that occurred in this area with some cases burning up to 70% of a municipality. Verdasca (2022) observed a decrease in the rate of expansion of the YLH in Portugal in 2018 and attributed this to the lack of biodiversity in recently burnt areas. In the northern region burnt areas weren’t even in the top 5 factors that may influence sightings despite being as prone to yearly wildfires as the centre.

Finally, “Shrublands”, a land use class characterized for having less than 10% tree cover and more than 25% of its area covered by some type of permanent vegetation was identified as having a strong negative effect on the density of sightings. Verdasca (2022) suggested these types of areas have the ideal conditions for predation (elevated flora and fauna biodiversity) however they lack the necessary structures for nest building.

2.2. Nest location vicinity

Carvalho, et al (2020) pointed out that YLH nests are more likely to be found on the canopies of trees due to the shelter these provide. However, reports of nests being found on the inside of walls, celling overhangs or other manmade structures are not unheard. As a matter of fact, the same author stated that anthropogenic structures were the second most likely place to find a nest. This trend remains the same with the updated data set. However the percentage of nests found on trees has decreased from 70.3% in 2020 to 67.27% in 2023 (table 10) while the percentage of nests found indoors increased from 6.1% to 7.59% in 2023. A slight increase of 0.41% in nests found in walls was also observed in this time frame.

Table 10 - Location Vespa velutina nests found and destroyed between 2013 and 2023

Structure	Proportion
Trees	67,27%
Indoor	7,59%
Roof	9,75%
Ground	3,00%
Walls	3,11%
Other	9,28%

In this study, “Artificial areas” are the main land use found around the areas where the nests were located, and the model reported a positive relationship between this predictor and the location of the nests. It is important to note, as stated by Johnston et al. (2020) that there is always a biased introduced with citizen science. In most cases, people will report nests that are either close to where they live or work or is part of their commute and they will avoid going out of their way to find YLH nests. This is one of the reasons “Artificial areas” is the main LU occupancy in the vicinity of these nests. Another reason is due to the abundance of food sources and shelter provided by urbanized areas.

“Temporary crops and pastures” were identified as being the second land use with the most influence on the YLH with a positive effect on the response. This may be in part because this land use class is characterized as being frequently managed by humans (increasing the likelihood of a nest being sighted). These areas of “Temporary crops and pastures” are typically small parcels of land adjacent to human construction used to plant annual crops. This not only makes identifying potential *V. velutina* nests more likely but also increases the number of pollinators in the vicinity, another source of protein for the hornets.

The combination of both “Permanent crops” and “Conifer forests” with a negative influence were also identified as being a combination of land uses less favourable to the YLH as the pure stands of conifers offer little biodiversity or populations of honeybees as observed on the landscape analysis at municipality level. The same is true for permanent crop plantations as combination of poor biodiversity and the constant use of phytochemicals make these areas unsuitable for the YLH as suggested by Otero et al (2015).

2.3. Impact on honeybees

It is commonly known that the Yellow Legged Hornet is a prolific hunter of honeybees and other wild pollinators (Verdasca, 2022). In its native Asian territory, it will prey on the native *A. cerana* which, through processes of coevolution have developed anti-predator behaviours. On the other hand, European honeybee (*A. melifera*) did not evolve these behaviours (Monceau et al., 2014) and therefore are far more susceptible to colony collapse because of excessive predation. Apiaries are an especially attractive source of protein because of the elevated concentration of honeybees in a relatively small area (Laurino et al., 2019).

The results in this study demonstrate a decline of honeybee colonies, and statistical evidence was found that this decline may be in part influenced by the density of hornets, ., on the North of Portugal. Between the years of 2018 and 2023 the decline was around 11-15%. This reflects the reality observed in the south of France where local beekeeper unions have reported losses of up to 7.5% between 2010 and 2011 (Monceau et al., 2014), 7 years after the first individuals were reported in the country. However, it is important to note that despite the decline in honeybee colonies, the overall decline in apiaries was not significant. This could be an indication that the impact caused by the YLH did not lead to abandonment from the beekeeper practices.

Leza et al. (2019) also studied the indirect impact that the YLH can have on local populations of honeybees and concluded that its presence can lead to adverse health impacts on them leading to diminished quality of life for the bees.

It is also important to point out that no evidence was found linking wildfires to the decline in honeybee populations, which could be another cause of apiaries and colonies decline in the North. However, the proportion of burnt areas in the North was reduced about 5%.

3. Final thoughts

While this study has helped shed light on many aspects of the biology and invasion patterns of the YLH, some improvements could be made. The land use classes used were far too broad, using specific species of forest stands or specific crops for the agricultural land could provide further information on the preference the YLH has for specific crops.

When looking at the artificial areas, the use of each area could also shed light on other patterns. Are industrial areas, where human presence is limited, more attractive in a nest building point of view? Or would they preferer residential areas where the diversity of tree cover and green areas may originate also diverse food sources? The effect that water can have on the dispersal, proximity to bodies of water versus flowing water. Finally, the effect suffered by honeybees should be revisited in a few years when the YLH has had more time to establish its presence.

References

Arca, M., Mougel, F., Guillemaud, T., Dupas, S., Rome, Q., Perrard, A., ... & Silvain, J. F. (2015). Reconstructing the invasion and the demographic history of the yellow-legged hornet, *Vespa velutina*, in Europe. *Biological Invasions*, *17*, 2357-2371.

Beggs, J. R., Brockerhoff, E. G., Corley, J. C., Kenis, M., Masciocchi, M., Muller, F., ... & Villemant, C. (2011). Ecological effects and management of invasive alien Vespidae. *BioControl*, *56*, 505-526.

Bessa, A. S., Carvalho, J., Gomes, A., & Santarem, F. (2016). Climate and land-use drivers of invasion: predicting the expansion of *Vespa velutina nigrithorax* into the Iberian Peninsula. *Insect Conservation and Diversity*, *9*(1), 27-37.

Budge, G. E., Hodgetts, J., Jones, E. P., Ostojá-Starzewski, J. C., Hall, J., Tomkies, V., ... & Stainton, K. (2017). The invasion, provenance and diversity of *Vespa velutina Lepeletier* (Hymenoptera: Vespidae) in Great Britain. *PLoS One*, *12*(9), e0185172.

Bunker, S. (2022). *Vespa velutina* (Asian hornet) [Dataset]. In CABI Compendium. https://doi.org/10.1079/cabicompendium.109164

Carvalho, J., Hipólito, D., Santarém, F., Martins, R., Gomes, A., Carmo, P., ... & Fonseca, C. (2020). Patterns of *Vespa velutina* invasion in Portugal using crowdsourced data. *Insect Conservation and Diversity*, *13*(5), 501-507.

Catarino, G. A. V. (2022). *Modelação Preditiva da Distribuição da Vespa Asiática (Vespa Velutina Nigrithorax): o caso de estudo das áreas urbanas funcionais de Coimbra e Viana do Castelo* (Doctoral dissertation).

Choi, M. B., Martin, S. J., & Lee, J. W. (2012). Distribution, spread, and impact of the invasive hornet *Vespa velutina* in South Korea. *Journal of Asia-Pacific Entomology*, *15*(3), 473-477.

Darrouzet, E., Gévar, J., Guignard, Q., & Aron, S. (2015). Production of early diploid males by European colonies of the invasive hornet *Vespa velutina nigrithorax*. *PLoS One*, *10*(9), e0136680.

Drake, J. A. (1989). Biological invasions: a global perspective. http://ci.nii.ac.jp/ncid/BA07272793

Franklin, D. N., Brown, M. A., Datta, S., Cuthbertson, A. G., Budge, G. E., & Keeling, M. J. (2017). Invasion dynamics of Asian hornet, *Vespa velutina* (Hymenoptera: Vespidae): a case study of a commune in south-west France. *Applied entomology and zoology*, *52*, 221-229.

ICNF, 2015. IFN6 – Principais resultados – relatório sumário [pdf], 34 pp, Instituto da Conservação da Natureza e das Florestas. Lisboa.

Johnston, A., Moran, N., Musgrove, A., Fink, D., & Baillie, S. R. (2020). Estimating species distributions from spatially biased citizen science data. *Ecological Modelling*, *422*, 108927.

Kenis, M., Rabitsch, W., Auger-Rozenberg, M. A., & Roques, A. (2007). How can alien species inventories and interception data help us prevent insect invasions?. *Bulletin of entomological research*, *97*(5), 489-502.

Kishi, S., & Goka, K. (2017). Review of the invasive yellow-legged hornet, *Vespa velutina nigrithorax* (Hymenoptera: Vespidae), in Japan and its possible chemical control. *Applied entomology and zoology*, *52*, 361-368.

Laurino, D., Lioy, S., Carisio, L., Manino, A., & Porporato, M. (2019). *Vespa velutina*: An alien driver of honeybee colony losses. *Diversity*, *12*(1), 5.

Leza, M., Herrera, C., Marques, A., Roca, P., Sastre-Serra, J., & Pons, D. G. (2019). The impact of the invasive species *Vespa velutina* on honeybees: A new approach based on oxidative stress. *Science of the Total Environment*, *689*, 709-715.

López, S., González, M., & Goldarazena, A. (2011). *Vespa velutina lepeletier,* 1836 (Hymenoptera: Vespidae): first records in Iberian Peninsula. *EPPO Bulletin*, *41*(3), 439-441.

Monceau, K., Arca, M., Leprêtre, L., Bonnard, O., Arnold, G., & Thiéry, D. (2018). How Apis mellifera behaves with its invasive hornet predator *Vespa velutina*?. *Journal of insect behavior*, *31*, 1-11.

Monceau, K., & Thiéry, D. (2017). *Vespa velutina* nest distribution at a local scale: An 8-year survey of the invasive honeybee predator. *Insect science*, *24*(4), 663-674.

Monceau, K., Bonnard, O., & Thiéry, D. (2014). *Vespa velutina*: a new invasive predator of honeybees in Europe. *Journal of pest science*, *87*(1), 1-16.

Norris, K. (2008). Agriculture and biodiversity conservation: opportunity knocks. *Conservation letters*, *1*(1), 2-11.

Galvão, J. (2021). *A Vespa velutina: um assunto de Proteção Civil?* (Doctoral dissertation).

Otero, I., Marull, J., Tello, E., Diana, G. L., Pons, M., Coll, F., & Boada, M. (2015). Land abandonment, landscape, and biodiversity: questioning the restorative character of the forest transition in the Mediterranean. *Ecology and Society*, *20*(2).

Quintans, S., 2020. Plano de ação para o controlo e vigilância da *Vespa velutina* em Portugal. PowerPoint presentation. DGAV

Requier, F., Rome, Q., Chiron, G., Decante, D., Marion, S., Menard, M., ... & Henry, M. (2019). Predation of the invasive Asian hornet affects foraging activity and survival probability of honey bees in Western Europe. *Journal of pest science*, *92*, 567-578.

Rome, Q., Muller, F. J., Touret-Alby, A., Darrouzet, E., Perrard, A., & Villemant, C. (2015). Caste differentiation and seasonal changes in *Vespa velutina* (Hym.: Vespidae) colonies in its introduced range. *Journal of Applied Entomology*, *139*(10), 771-782.

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., ... & Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nature communications*, *8*(1), 14435.

Tan, K., Radloff, S. E., Li, J. J., Hepburn, H. R., Yang, M. X., Zhang, L. J., & Neumann, P. (2007). Bee-hawking by the wasp, *Vespa velutina*, on the honeybees Apis cerana and A. mellifera. *Naturwissenschaften*, *94*, 469-472.

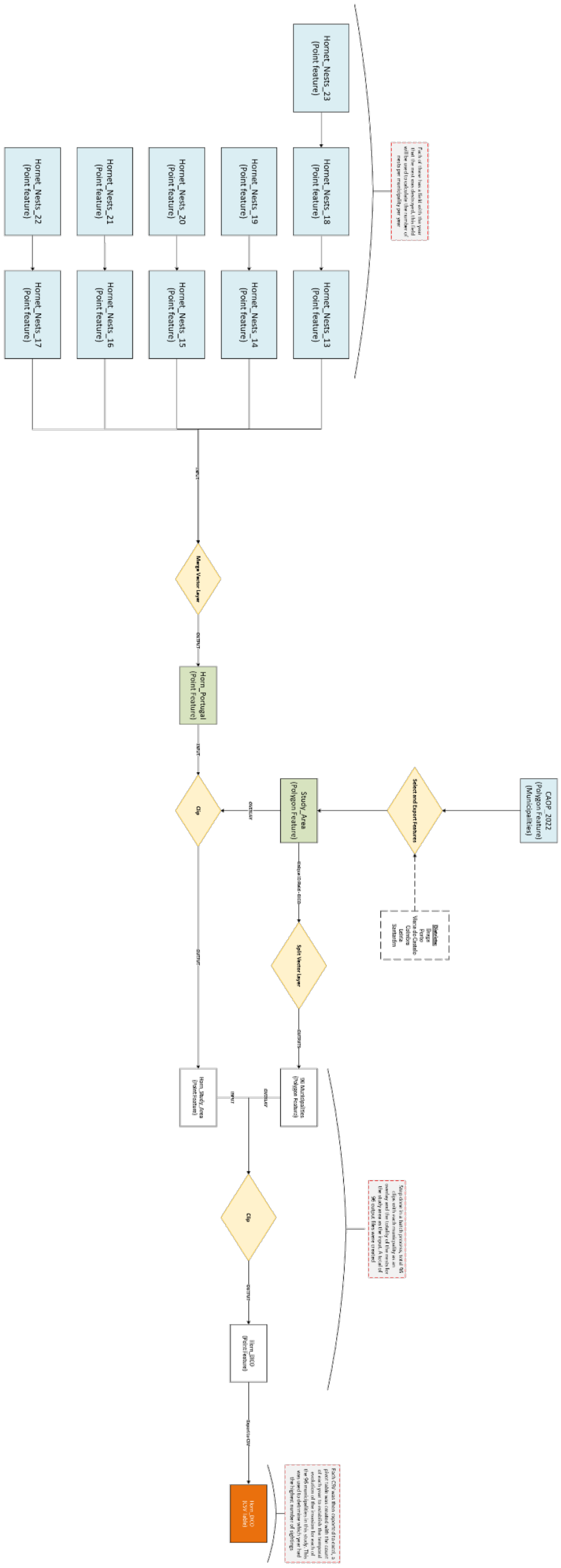
Verdasca, M. J. (2022). Evaluation of the invasion dynamics of non-native species and associated socio-ecological impacts: the case study of *Vespa velutina*.

Villemant, C., Streito, J. C., & Haxaire, J. (2006). Premier bilan de l'invasion de *Vespa velutina Lepeletier* en France (Hymenoptera, Vespidae). *Bulletin de la Société entomologique de France*, *111*(4), 535-538.

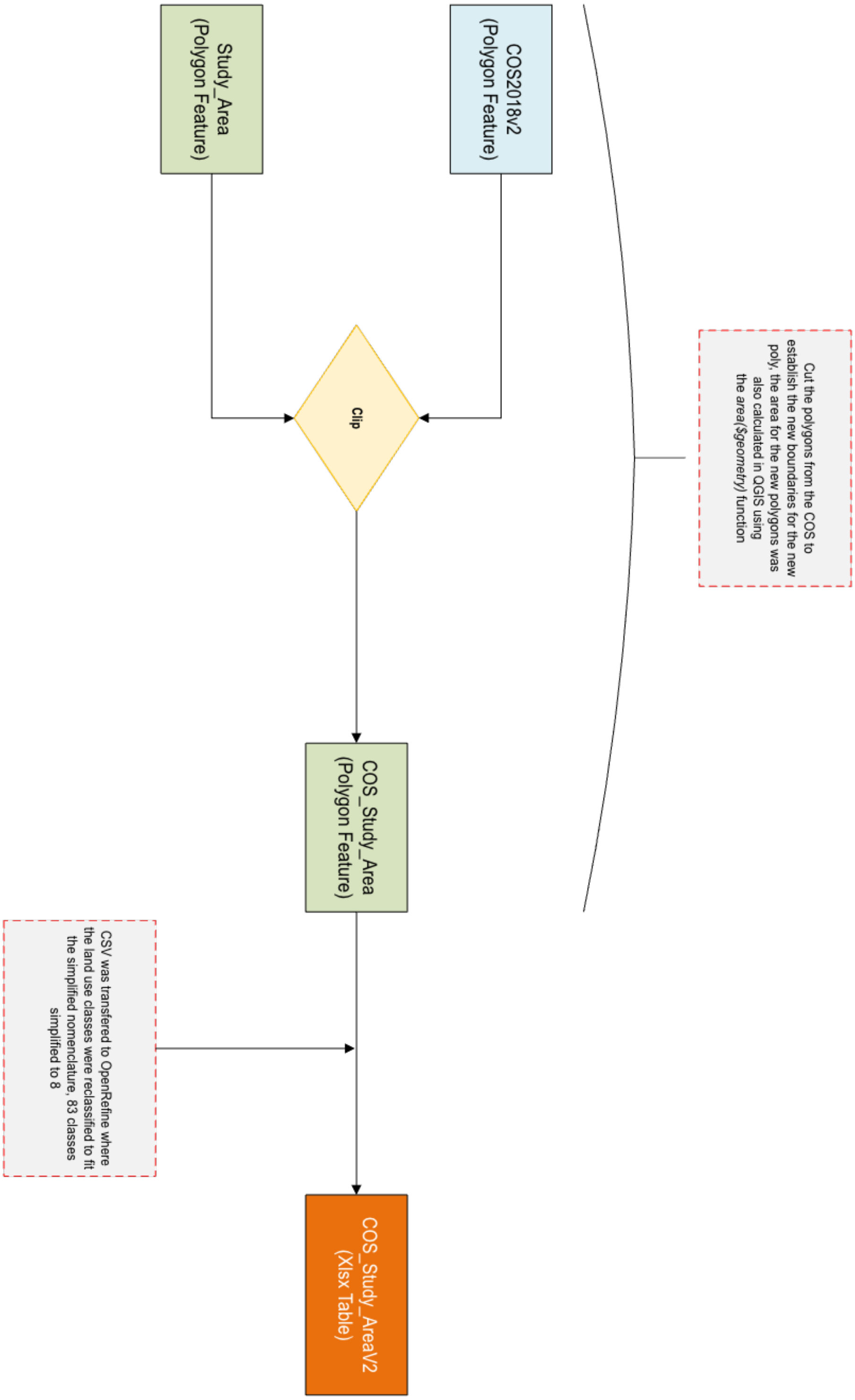
Walter, J., Görner, T., Šulda, L., Bureš, J., Myslík, Z., Milička, R., ... & Brus, J. (2024). First Czech record of the Asian hornet (*Vespa velutina*) and a climatic prediction of its spread in the Czech Republic.

Annex

Annex 1 – Flowchart for the projection and selection of the hornet nests post curation



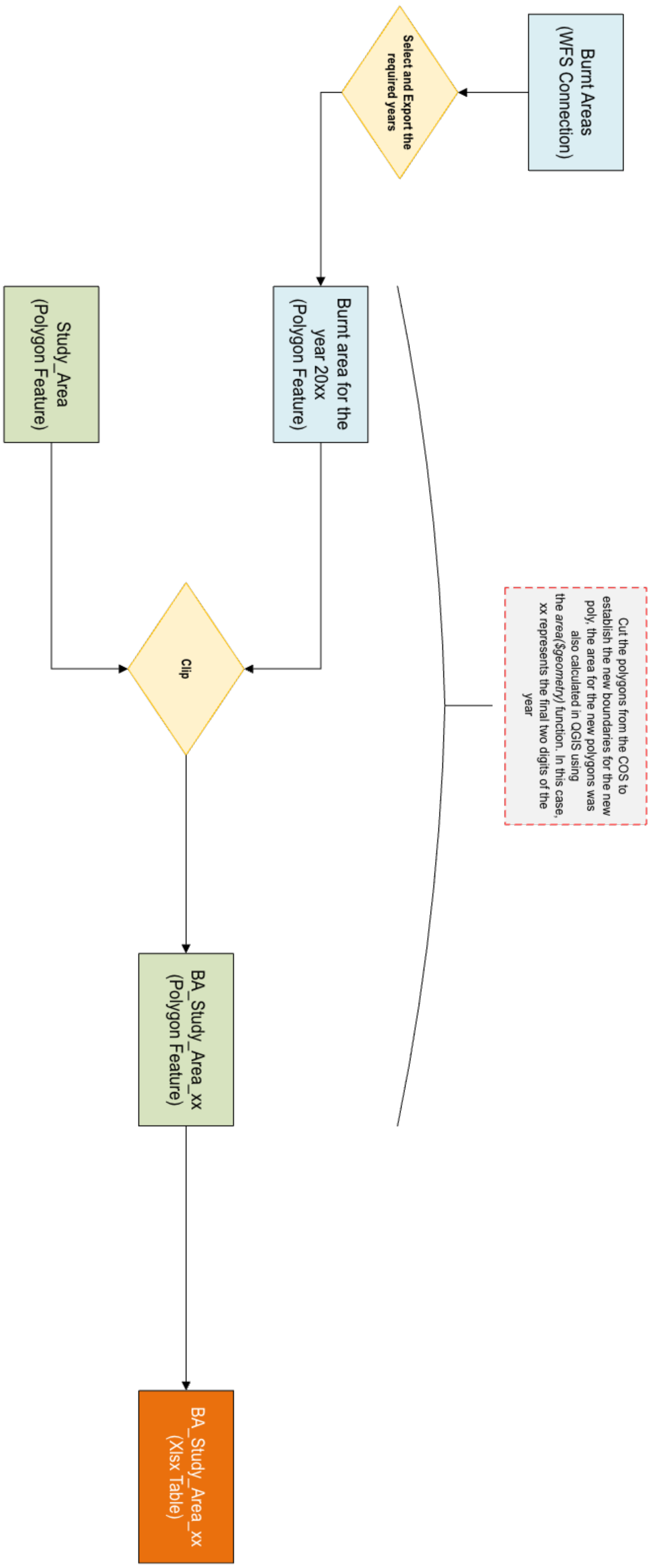
Annex 2 – Flowchart for the intersection of the land use data with the study area



Cut the polygons from the COS to establish the new boundaries for the new poly, the area for the new polygons was also calculated in QGIS using the *area($geometry)* function

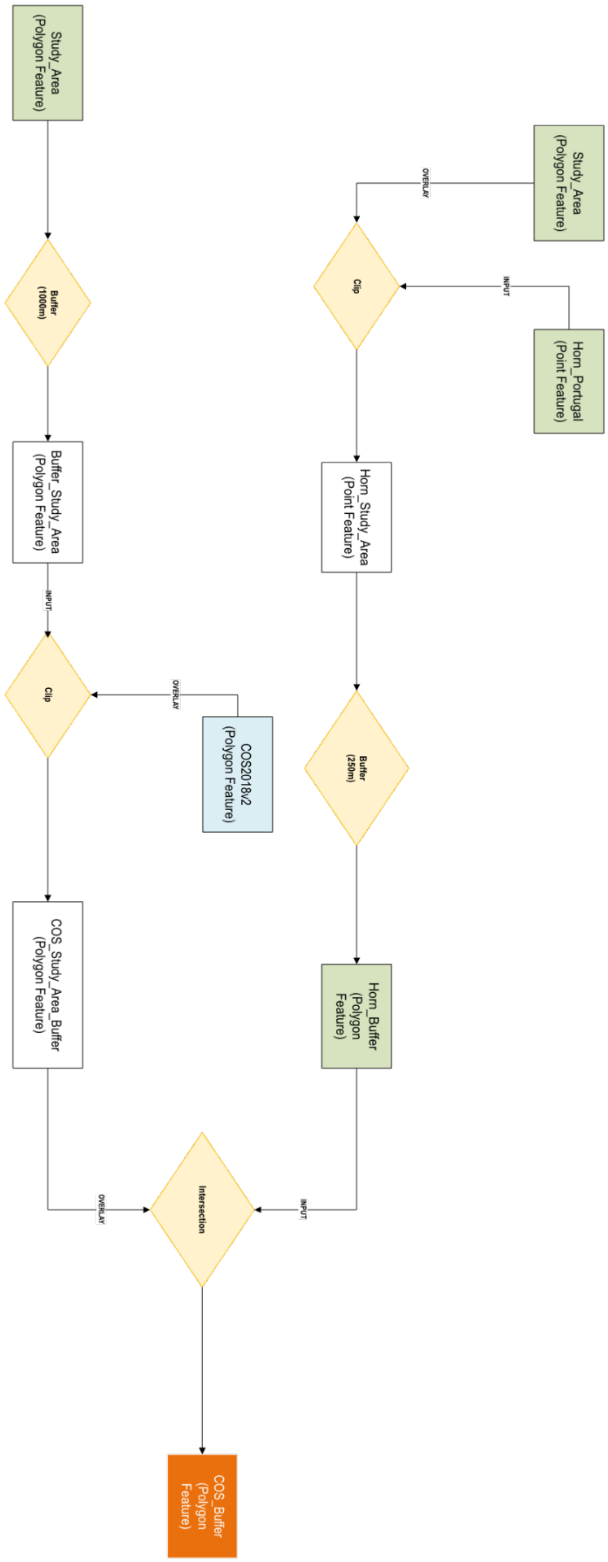
CSV was transfered to OpenRefine where the land use classes were reclassified to fit the simplified nomenclature, 83 classes simplified to 8

Annex 3 – Flowchart for the intersection of the burnt area with the study area

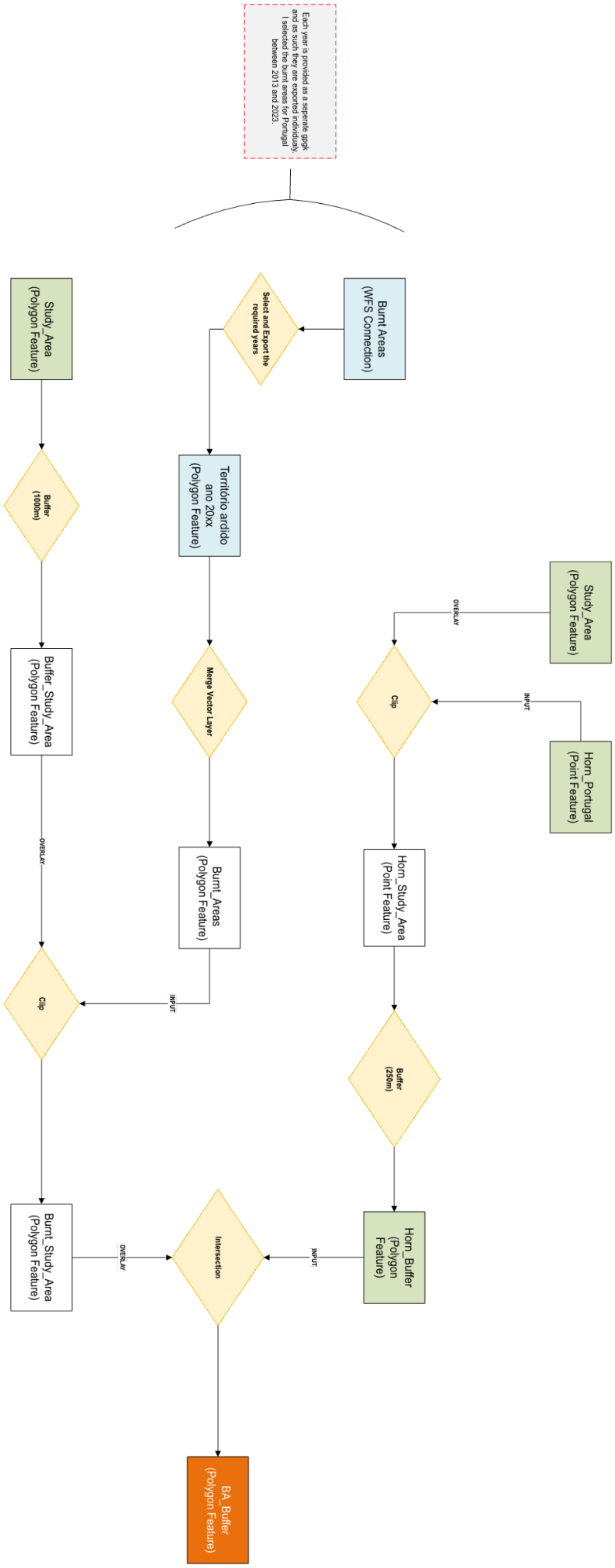


Cut the polygons from the COS to establish the new boundaries for the new poly, the area for the new polygons was also calculated in QGIS using the *area($geometry)* function. In this case, xx represents the final two digits of the year

Annex 4 – Flowchart for the intersection of the land use data with the hornets' activity buffer



Annex 5 – Flowchart for the intersection of the burnt area with the hornets' activity buffer



Annex 6 – Illustration of the proportion of land use for the centre and northern region as well as globally for the 96 municipalities