Faculty of Science
Department Biology

Wild boar (*Sus scrofa* L.) distribution and agricultural damage in Flanders

Verspreiding van het everzwijn en landbouwschade in Vlaanderen

English

Thesis submitted for the degree of doctor in Biology at the University of Antwerp to be defended by

Anneleen RUTTEN

Promotors:

Prof. dr. Herwig Leirs

Ir. Dr. Jim Casaer

Antwerp, 2019
Internal PhD committee

Prof. dr. Erik Matthysen
Prof. dr. Stijn Temmerman
Prof. dr. Rudy van Diggelen

External PhD committee

Prof. dr. Herbert Prins (Wageningen University, Wageningen, the Netherlands)
Dr. Oliver Keuling (University of Veterinary Medicine Hannover, Hannover, Germany)
“You are capable of more than you know. Choose a goal that seems right for you and strive to be the best, however hard the path. Aim high. Behave honorably. Prepare to be alone at times, and to endure failure. Persist! The world needs all you can give.”

Edward O. Wilson
Table of Contents
Summary

1

Samenvatting

7

Chapter 1

Introduction

Chapter 2

Using an online survey to assess the spatial distribution of wild boar crop damage and factors influencing this distribution

Chapter 3

Assessing agricultural damage by wild boar using drones

Chapter 4

What attracts wild boar to agricultural crops?

Chapter 5

Analysing the recolonization of a highly fragmented landscape by wild boar using a landscape genetic approach

Chapter 6

Future distribution of wild boar in a highly anthropogenic landscape: models combining hunting bag and citizen science data

Chapter 7

Discussion

Dankwoord

169

References

177

Supplementary materials

207
Summary
The Eurasian wild boar (*Sus scrofa* L.) is one of the most widespread mammals of the world. Wild boar are characterised by their opportunistic omnivorous behaviour, high reproductive potential, flexible habitat selection and high adjustment potential to anthropogenic disturbances. This caused the species to expand in both their native as well as their invasive distribution range. Wild boar can have a considerable environmental impact and are therefore often considered as landscape engineers. However, due to population expansions and landscapes which become increasingly anthropogenic and fragmented, wild boar increasingly come into contact with human activities causing Human-Wildlife Impacts. In Flanders (northern Belgium), wild boar re-emerged in 2006 after almost half a century of absence due to local extinction. During their absence, the Flemish landscape changed dramatically due to economic growth, urbanization and agricultural intensification. This results in Flanders currently being one of the most densely human populated regions in Europe characterised by a severely fragmented and anthropogenic landscape.

Wild boar is known to show a substantial plasticity in adjusting towards human-dominated environments. Currently, the Flemish wild boar distribution is limited to the East of Flanders with the largest distribution range in the province of Limburg and increasing distribution ranges in the provinces of Antwerp and Vlaams Brabant. With wild boar populations currently increasing in Flanders, the Flemish situation presented an interesting case study to assess if there are limits towards the flexibility of wild boar in adjusting to anthropogenic pressures. Moreover, due to economic costs of wild boar impacts, stakeholder acceptance of wild boar being part of the Flemish wildlife is challenged. Especially crop damage is raising concerns because this damage is not monitored in Flanders and the extent is therefore unknown. By getting a better understanding of current and potential crop damage distribution and the effect of fragmentation on damage risk, the extent of crop damage as a Human-Wildlife Impact can become clear. This can lead to a more efficient implementation of management tools to prevent these Human-Wildlife Impacts.
An online survey, sent to farmers in Limburg, gave us a first documentation of the geographic distribution of crop damage. The survey gave us the opportunity to select a study area in which most damage was reported. This study area was used to record damaged fields between 2015 and 2018. By the use of a drone, damaged fields were photographed. Based on Geographic Object-Based Image Analysis and Random Forest models, we developed a time-efficient, standardized and accurate method to assess the damaged area within a field with high resolution. The dataset on damaged fields also allowed us to get a better understanding of factors attracting wild boar to a specific field. By assessing both the landscape around the field taking into account habitat fragmentation, as well as field-specific characteristics, we could make predictive models based on these characteristics to assess damage risk throughout Flanders. Moreover, the gained information can now help taking preventive measures more efficiently or to guide agro-technical adjustments to optimally limit damage risk.

To get a better insight in factors driving recolonization and expansion, a landscape genetic research was conducted. By genotyping DNA-samples from hunting bag, we showed that the wild boar population expansion started from 2 local gene pools in the province of Limburg. Since 2015 a third gene pool reaching Flanders through natural migration from the Netherlands emerged in the Northwest of Flanders. The landscape genetic analysis revealed the importance of forest in explaining spatial genetic patterns while fragmentation was found not to have an influence on wild boar distribution in this highly fragmented landscape.

Moreover, a species distribution model was set up to assess remaining habitat suitability in Flanders allowing us to get a better understanding of the future expansion of wild boar in Flanders. With forest being the main factor determining habitat suitability, expansion of the Flemish wild boar population can be expected to remaining areas with forests in the severely fragmented landscape of Flanders. With its immense plasticity, wild boar is not affected by fragmentation of their habitat even in the most severely fragmented and most
densely populated landscapes of Europe. Moreover, the behavioural plasticity of wild boar in adjusting towards anthropogenic pressures may not yet have reached its limits which could mean that with future dispersion, wild boar may also colonise new types of habitat.

Gaining this knowledge was essential to conduct a risk assessment with the current and future distribution of wild boar in Flanders. With a better understanding on the current and future extent of crop damage causing Human-Wildlife Impacts, management actions and preventive measures can be taken more efficiently.
Samenvatting
Samenvatting
Everzwijnen behoren tot de meest wijdst verspreide zoogdieren van de wereld. Everzwijnen worden getypeerd door hun opportunistisch gedrag, hoge reproductiviteit, flexibiliteit qua habitat selectie en een hoog aanpassingsvermogen aan menselijke aanwezigheid. Hierdoor nemen hun populaties toe, zowel binnen hun inheemse verspreidingsareaal, als in gebieden waar ze uitheems en invasief zijn. Everzwijnen staan bekend als landschapsingenieurs omdat ze een behoorlijk grote impact kunnen hebben op de natuur. Echter, door de toenemingen in aantallen in combinatie met de verstedelijking van landschappen, komen everzwijnen ook meer en meer in contact met menselijke activiteiten. Zo ook in Vlaanderen waar het everzwijn sinds 2006, na een halve eeuw afwezigheid, terug is van weg geweest. Tijdens hun afwezigheid is het Vlaamse landschap behoorlijk veranderd, onder meer door de ontwikkeling van intensievere landbouw, economische groei en urbanisatie. Op dit moment is Vlaanderen een van de dichtst bevolkte regio’s van Europa, met een sterk gefragmenteerd en antropogeen landschap. Bosfragmenten en landbouwgebieden zijn hier zodanig nauw met elkaar verweven dat contact tussen everzwijnen en menselijke activiteiten onvermijdbaar

Momenteel is het everzwijn vooral te vinden in Limburg en is zijn verspreidingsgebied in Antwerpen en Vlaams-Brabant aan het toenemen. De stijgende populaties van everzwijnen in Vlaanderen vormen een interessante casus om te onderzoeken of er limieten zijn aan de flexibiliteit van everzwijnen om zich aan te passen aan sterk verstedelijkte gebieden. Door economische kosten die gepaard gaan met de impact van everzwijnen, wordt de acceptatie van everzwijnen als deel van onze natuur op de proef gesteld. Vooral landbouwschade krijgt meer en meer aandacht omdat de omvang ervan niet wordt opgevolgd in Vlaanderen en we dus niet weten hoeveel schade er in werkelijkheid is. Om die reden is het essentieel een beter inzicht te verkrijgen in de omvang van landbouwschade.

Een online enquête die verstuurd werd naar landbouwers in Limburg gaf ons een eerste beeld van de verspreiding van landbouwschade. De enquête bood ons
Samenvatting

de mogelijkheid om een studiegebied te selecteren op basis van de meeste schademeldingen. In dit Limburgse studiegebied werden alle gemelde schadegevallen tussen 2015 en 2018 opgevolgd. Aan de hand van een drone werden de beschadigde velden gefotografeerd en via Geographic Object-Based Image Analysis en Random Forest modellen kon, aan de hand van de spectrale informatie van deze foto’s, bepaald worden wat het beschadigde aandeel van een veld was. Dit resulteerde in een tijdsefficiënte, gestandaardiseerde en nauwkeurige methode die nu kan worden ingezet om de beschadigde oppervlakte van een perceel gedetailleerd te bepalen. Verder liet deze dataset van beschadigde percelen liet ons ook toe om een beter inzicht te krijgen in factoren die everzwijnen naar een bepaald perceel aantrekken. Door zowel het landschap rondom het veld te analyseren, rekening houdend met fragmentatie, als door te kijken naar perceel-specifieke factoren zoals bemesting, maisvariëteit etc. was het mogelijk om voorspellingsmodellen maken om het schaderisico in heel Vlaanderen op individueel perceel-niveau te bepalen. Dit laat toe om gerichter preventieve maatregelen te nemen of landbouw-technische aanpassingen te implementeren om schade zo veel mogelijk te beperken.

Om een beter inzicht te krijgen in de factoren die rekolonisatie en verspreiding van everzwijnen beïnvloeden, werd een landchaps-genetische analyse uitgevoerd. Via genotypering van DNA-stalen afkomstig van het afschot van everzwijnen in Vlaanderen konden we bepalen dat de rekolonisatie startte vanuit 2 lokale genenpoelen. Vanaf 2015 is er een derde genenpoel terug te vinden in het noorden van Vlaanderen die verbonden is met de Nederlandse populatie. De landschaps-genetische analyse kon het belang van bos als drijver voor de gevonden spatiale genetische patronen aantonen. Deze analyse toonde bovendien aan dat fragmentatie geen invloed heeft op de verspreiding van het everzijn.

Tenslotte werd een verspreidingsmodel opgesteld om overblijvend geschikt habitat in de rest van Vlaanderen te bepalen. Aan de hand hiervan werd een beeld verkregen van de mogelijke toekomstige verspreiding van everzwijnen. In de eerste plaats verwachten we toekomstige verspreiding in bosrijke gebieden,
Samenvatting

aangezien bos de meest bepalende factor is voor geschikt habitat. Ten tweede worden ze door hun enorme plasticiteit zelfs in dit sterk gefragmenteerde en dicht bevolkte landschap van Europa niet beïnvloed door habitatfragmentatie. Als laatste kunnen we stellen dat everzwijnen de grens van hun aanpassingsvermogen naar alle waarschijnlijkheid nog niet hebben bereikt waardoor ze bij hun verdere verspreiding mogelijks nieuwe habitattypes van zouden kunnen koloniseren.

Met de kennis die in dit onderzoek werd vergaard hebben we aspecten die essentieel zijn voor het uitvoeren van een risicoanalyse van de huidige en toekomstige verspreiding van everzwijnen in Vlaanderen verduidelijkt. Deze kennis over de huidige en toekomstige omvang van landbouwschade, de grote bron van impact op menselijk activiteiten, laat nu toe om beheersmaatregelen en preventieve maatregelen efficiënter toe te passen.
Samenvatting
Chapter 1

Introduction
The Eurasian wild boar (*Sus scrofa*, Linnaeus, 1758) is a native species in Flanders (northern Belgium) but they locally got extinct for more than half a century. Although the current Flemish landscape is characterized by a severely fragmented and urbanized landscape, wild boar returned in 2006. Since their recolonization, populations keep increasing and their distribution range extents. Wild boar has been shown to have a substantial behavioural plasticity in adjusting to human-dominated environments: wild boar tolerate human presence by modulating their risk perception and have even been shown to use human associated habitat classes (Cahill et al. 2012, Stillfried et al. 2017). However, it has been shown that landscape fragmentation affects wild boar occurrences and they adjust their spatio-temporal behaviour in fragmented landscapes (Virgós 2002, Podgórski et al. 2013). At the moment it is unknown if there is a limit to the behavioural plasticity of wild boar thereby impeding the further colonization of human-dominated environments and fragmented landscapes. Moreover, due to close contact with human activities, wild boar presence causes challenges due to impacts on human activities. There is no scientific information on the extent of these impacts nor on implementation of measures to prevent or limit these impacts in such landscapes. Flanders, being one of the most densely populated areas of Europe, characterized by a severely fragmented landscape, represents a very extreme example of this kind of highly fragmented, human dominated landscapes and therefore offers an ideal case study to examine these knowledge gaps. Therefore, the aim of this PhD thesis is to provide a sound scientific basis of lacking knowledge to be able to conduct a risk assessment of wild boar impacts under current conditions and with their potential future dispersal in Flanders.

Wild boar

The Eurasian wild boar is an ungulate species part of the Suidae family, in the order Artiodactyla. Wild boar are large mammals characterized by a strong muscled body, brown to greyish colour, typical long snout and pronounced canines in adult males. Body size and weight of adult wild boar show large
regional differences with shoulder heights between 60 and 110 cm and weights which can go higher than 150 kg (Casaer and Van Den Berge 2006). Sexual dimorphism results in females generally being smaller than males (Pedone et al. 2014). Piglets have typical longitudinal light coloured stripes on their back which have a role in camouflage (Fang et al. 2009). With the highest reproductive potential among ungulates, wild boar have litters ranging between 3 and 12 (Frauendorf et al. 2016, Rosell et al. 2012) after a gestation period of 115 days (Henry 1968). A well-developed smell and auditory capacities allow wild boar to forage, communicate and observe their surroundings, although their visual capacities are not well developed (Keuling et al. 2018).

The social structure of wild boar is characterized by maternal groups composed by a few adult females, their piglets and yearlings (sub adults) while adult males are solitary (Dardaillon 1988, Maselli et al. 2014). Variation can be found in group compositions and group structures can change due to disruption (Fernández-Llario et al. 1996, Iacolina et al. 2009). Wild boar are not territorial but are typically characterized by a home range which varies depending on sex, food availability and population density (Boitani et al. 1994, Massei et al. 1997c). Wild boar movements are mainly driven by the search for food, dispersal and fleeing after disturbance (Keuling et al. 2008b, Massei et al. 1997a, Morelle and Lejeune 2015). In a natural undisturbed area, wild boar are usually diurnal (Keuling et al. 2008c, Kurz and Marchinton 1972, Ohashi et al. 2013). However, due to human disturbance or changes in population density, there can be a shift to a more nocturnal activity pattern (Keuling et al. 2008c, Massei et al. 1997a, Ohashi et al. 2013).

The wild boar is an opportunistic omnivorous species. Depending on food availability, seasonality and habitat variability, their diet can vary substantially (Massei and Genov 2004, Schley and Roper 2003). Plant matter mostly represents the largest part of their diet and is composed by mast seeds (acorns, chestnuts and beech), roots, fruits, and agricultural crops (Genov 1981, Massei and Genov 2004, Schley and Roper 2003). Depending on availability, agricultural crops can be an important part of consumption (Herrero et al.
The importance of animal matter in a boar’s diet such as earthworms, grubs and other macro-invertebrates, eggs, vertebrate remains etc. varies seasonally (Dardaillon 1987) and can be an important nutritional addition to their diet (Ballari and Barrios-García 2014).

Wild boar can occupy a large variety of habitats which are mainly defined by food availability and shelter possibilities. Typical habitats therefore include forests, scrublands and agricultural areas (Briedermann 1990, Keuling et al. 2018). Habitat utilization by wild boar have been shown to be adaptive and highly flexible in response to anthropogenic disturbances such as hunting (Calenge et al. 2002, Keuling et al. 2008c, Maillard and Fournier 2014). Moreover, variation in habitat selection occurs throughout the year because of changing food availability (Keuling et al. 2008a, Thurfjell et al. 2009).

Population densities are mainly defined by food availability and temperature (Geisser and Reyer 2005, Melis et al. 2006). Favourable climatic conditions reduce juvenile mortality as piglets have a poor thermo-insulative ability making frost temperatures difficult to survive (Graves 1984). Food availability for wild boar varies with mast years, the intensification of agriculture and supplementary feeding. As a typical pulse-resource species, changing food availability affects wild boar reproductive success in different ways. 1- with increasing food availability litter sizes increase and 2- piglets build up weight fast to survive cold temperatures (Bruinderink et al. 1994, Fernández-Llario et al. 1999, Geisser and Reyer 2005). 3- As wild boar fecundity is defined by a certain body mass threshold between 25 and 35 kg, also the age of first reproduction decreases with increasing food availability (Gaillard et al. 1993, Sabrina et al. 2009). Wild boar are predated by wolves, tigers, dingoes and bobcats (Forsyth et al. 2019, Hayward et al. 2012, Nores et al. 2008, Stegeman 1938). However, wolves seem to have only limited impact on wild boar populations (Głowaciński and Profus 1997, Melis et al. 2006, Nores et al. 2008). Hunting is often reported to be the main cause of wild boar mortality, however often insufficient to limit wild boar population growth (Massei et al. 2015, Nores et al. 2008).
Wild boar can have a considerable environmental impact through feeding, rooting and predation and are therefore often considered landscape engineers or ecological keystone species (Barrios-Garcia and Ballari 2012, Meijaard and Melletti 2018). By their rooting behaviour, they can have positive as well as negative impacts on plant communities as the dynamics of rooting varies with habitat and soil types (Welander 2000). The nature of ecological impacts varies across habitat types and ecosystem types. Positive effects of wild boar on the environment include reductions of competitive plants, the role they play in seed dispersal and increasing plant species diversity in grasslands due to rooting (Horčičková et al. 2019, Massei and Genov 2004). Negative effects include decline of herbaceous understory in forests and changes in plant species composition (Bratton 1974). Also impacts on animal communities range from negative impact on ground-nesting birds due to predation on their eggs to a significant decrease of damage by insect pests in tree plantations due to predation on these insects (Massei and Genov 2004).

**Human-Wildlife Impacts**

*Expansion capacity of wild boar*

The wild boar is a species with a high behavioural flexibility, they can adjust to anthropogenic disturbances and have a high reproductive potential (Massei et al. 2015, Rosell et al. 2012). These characteristics cause worldwide expansions of wild boar populations (Figure 1.1). Since the 1960s, wild boar populations have been expanding throughout Europe (Acevedo et al. 2007, Massei et al. 2015, Saez-Royuela and Telleria 1986). A combination of environmental, agricultural and demographic changes have been reported to trigger these expansions: 1- climate change leads to higher winter temperatures favouring reproductive success and juvenile survival, 2- agricultural intensification with increasing maize cultivation, 3- supplementary feeding and 4- low densities (or even a lack) of predators (Massei et al. 2015, Morelle et al. 2016, Pittiglio et al. 2018, Wotschikowsky 2010).
Pigs and wild boar were an important food source for explorers taking voyages to America in the 15\textsuperscript{th} century or Australia in the 18\textsuperscript{th} century where they were deliberately released to provide a food source for future explorers and to provide a new game species for hunting (Choquenot et al. 1996, Mayer 2018, Meijaard and Melletti 2018). Moreover, accidental escapes of domestic pigs and interbreeding with wild boar resulted in feral pigs (Mayer 2018). Because of their flexibility and high reproductive success, also in their invasive distribution range populations are expanding (New Zealand: Bengsen et al. 2018, Australia: Choquenot et al. 1996, North America: Mayer 2018, & McClure et al. 2015, South America: Salvador and Fernandez 2018). This makes the wild boar currently one of the most widespread mammals of the world found in all continents apart from Antarctica (Barrios-Garcia and Ballari 2012, Keuling et al. 2018).

\textbf{Figure 1.1:} Worldwide distribution of wild boar with their native distribution range in black and introduced range in grey. Circles represent introductions on islands, question marks represent occurrence with unknown distribution. From Barrios-Garcia and Ballari (2012).
Chapter 1

Sources of conflicts

The downside of these wild boar population expansions in more anthropogenic areas is the increasing number of contacts with human activities causing conflicts. Human-Wildlife Conflicts causing a high economic impact results in wild boar sometimes regarded as a pest- or invasive species (Bieber and Ruf 2005, Morelle et al. 2016). Human-Wildlife Conflicts because of wild boar presence include traffic collisions, disease transmission and damage to agricultural crops.

Although road mortality of most wild boar populations contribute only marginally to the total mortality of wild boar, the expansion of our road network fragments natural habitats causing traffic safety hazards and economic losses (Bruinderink and Hazebroek 1996, Primi et al. 2009). No numbers on vehicle collisions are available for Flanders. However 30 traffic accidents per 1000 km² per year with wild boar are reported in the Northwest of Spain, an area with a road density almost ten times lower than in Flanders (0.59 km/km² versus 5.08 km²) (Lagos et al. 2012) and 12 traffic accidents per 1000 km² per year for central Italy in an area with a road density of 0.43 km/km² (Primi et al. 2009). In Sweden, about 1000 vehicle collisions wild boar occur per year, in the Netherlands 320 collisions with wild boar occurred in 2012 and in Germany 13700 traffic accidents involved wild boar in 2005 (Liberg et al. 2010, Van Wieren and Groot Bruinderink 2010, Wotschikowsky 2010).

As wild boar and domestic pigs share common pathogens and wild boar can act as a disease reservoir, the economic impact of disease outbreaks in domestic pigs originating from wild boar populations can be large (Cvetnic et al. 2003, Ruiz-Fons et al. 2006). Diseases specific in wild boar include classical swine fever, African swine fever, Aujeszky’s disease (pseudo rabies virus), brucellosis, tuberculosis, tick borne encephalitis, trichinellosis, leptospirosis, hepatitis E etc. (Meng et al. 2009, Ruiz-Fons et al. 2008, Ruiz-Fons 2017). Some of these diseases have high prevalence in European wild boar populations and are actively monitored in Flanders by the Flemish Nature and Forestry Agency (ANB). None of these diseases currently pose a risk in Flanders.
However since September 2018, an outbreak of African swine fever occurred in the South of Belgium (Scicom 2018). Given the economic implications on domestic pig industry by prohibiting export of meat once the disease enters an area, a set of preventive measures have been taken to prevent further spread of the disease in the wild boar population and avoid outbreaks in domestic pigs within the core area, buffer zones as well as in the rest of Belgium (Scicom 2019). At this moment, no further outbreaks have been found.

Moreover, damage to agricultural crops can cause severe economic losses for farmers (Massei et al. 2015, Schley et al. 2008). The yearly damage compensations paid due to crop damage reflect the potential economic impact of crop damage by wild boar for individual farmers. In the Netherlands, where damage is reported to the Faunafonds, the annually paid refunds for crop damage ranged between €97 000 and €190 000 over the period 2009 until 2014 (Faunafonds 2014). In Wallonia compensations paid yearly varied between €237 000 and €720 000 between 2008 and 2014 (Widar and Luxen 2016). In France, wild boar damage compensations increased from less than one million euro in 1970 to more than €20 million annually since 2000 (Carnis and Facchini 2012). In Luxembourg, yearly damage compensations increased from €100 000 in 1971 to €900 000 in 2004 (Schley et al. 2008).

From conflicts to impacts

Human-Wildlife Conflicts are called conflicts because of the economic or social impact they cause. However framing these impacts as conflicts, presenting wildlife as antagonists may alter the way people perceive those species and has a negative connotation often limiting incentives for finding effective solutions (Redpath et al. 2015, Young et al. 2010). By adjusting this term to Human-Wildlife Impacts, defined as impacts due to interactions between wildlife and humans activities (including both impacts of wildlife on humans as well as impacts of humans on wildlife) according to Redpath et al. (2013), this presents the opportunity to understand the origin of these impacts, focus on solutions,
involve underlying human dimensions and stakeholder concerns and interests (Frank 2012).

The importance of Human-Wildlife Impacts leads to studies conducted to increase understanding on spatial and temporal patterns and causes of these impacts, in order to develop appropriate preventive measures. For traffic collisions this includes combinations of fencing and wildlife passages or light warning signs which are triggered by wildlife (Bruinderink and Hazebroek 1996, Lagos et al. 2012, Primi et al. 2009, Rodríguez-Morales et al. 2013). Research towards the importance of wild boar as a disease reservoir for example resulted in a ban to translocate wild boar for hunting purposes in Spain (Ruiz-Fons et al. 2007). Finally, assessing spatio-temporal characteristics of crop damage can contribute to the development of a more effective management planning (Lombardini et al. 2016).

Essential in the assessment of Human-Wildlife Impacts are underlying human dimensions, stakeholder concerns and interests (Frank 2012). As the economic effect such as crop damage by wild boar is the main limiting factor in stakeholders’ tolerance towards wildlife, the key to successful wildlife management is to include both ecological and human dimensions (Carpenter et al. 2013, Keuling et al. 2016). By considering human dimensions in the decision making process of wildlife management, also positive impacts can be emphasized which are valued by all stakeholders (Riley et al. 2003). Positive impacts include the recreational value for people spotting wild boar therefore, being an added value for ecotourism, recreational hunting, meat production, the recognition of wild boar as part of (restoring) ecosystem functioning and finally wild boar has been reported to act as a replacement hunting target in Brazil thus preventing overhunting of native wildlife (Barrios-Garcia and Ballari 2012, Desbiez et al. 2011).
Wild boar in Flanders

Return of wild boar to Flanders

Flanders is located in the Northeast of Europe and is the northern part of Belgium. Wild boar have always been present in neighbouring regions of Flanders like Wallonia (southern Belgium) Germany, Luxembourg and France; and after local extinction around 1830, wild boar was reintroduced in the Netherlands for hunting in 1907 in fenced areas from which regular escapes occurred (Casaer and Licoppe 2010, Guldemond et al. 2015, Massei et al. 2015, Van Wieren and Groot Bruinderink 2010). In line with European population expansions since the 1960’s, wild boar populations in these regions increased the last decennia (Massei et al. 2015, Saez-Royuela and Telleria 1986). In Flanders, wild boar got locally extinct after the Second World War due to overhunting. In 2006, after more than half a century, wild boar returned to Flanders (Scheppers et al. 2014). The recolonization started from two geographically distinct locations in the eastern province of Limburg. These recolonization sites were geographically not connected to neighbouring populations excluding a return by recolonization through natural migration. This implies that the recolonization probably started through local (illegal) reintroductions or escapes. However, there is no official information on where these populations originate from as well as there is no registration of reintroductions nor escapes.

Fragmented landscape

As in many parts of western Europe, the Flemish landscape changed dramatically over the last century after World War II due to an economic boom. Existing landscapes changed into a completely new environment as a result of urbanization, agricultural intensification and increasing human pressure on rural landscapes (Antrop 1997). Flanders is densely populated and highly urbanized. Flanders has 462 inhabitants/km² while Wallonia only has an average
population density of less than 50 inhabitants/km². This high population density makes Flanders one of the most densely populated areas in Europe (Antrop 2004, FOD Economie 2011, Linell et al. 2001). Build-up area increased with 39% between 1980 and 1992, major roads and railways occupy 7.6% of the land surface resulting in a road network of 5.08 km/km², 53% of the land cover is used for agriculture and only 11.4% is covered by forest (Antrop 1997, 2004, Demolder et al. 2014, Vercayie and Herremans 2015).

Next to nutrient enrichment, soil dehydration, climate change and invasive species, fragmentation and loss of habitat are mentioned as having important negative impacts on the biodiversity in Flanders (Demolder et al. 2014). Nonetheless, due to new conservation strategies, protective legislation, increasing stakeholder acceptance and supportive public opinion (Chapron et al. 2014, Pereira and Navarro 2015, Trouwborst 2010), several large mammal species which were previously extinct are returning to Flanders including beavers, otters, and recently also wolves (Swinnen and Rutten et al. 2019). Also wild boar, with their flexible characteristics, has managed to start its recolonization in this highly fragmented landscape.

**Hunting legislation in Flanders**

According to the hunting legislation in Flanders, wild boar is regarded as a native game species. Hunting rights belong to the landowner who can lease the hunting rights, hunt themself or decide not to allow hunting (Casaer and Licoppe 2010). Hunting rights on an area of at least 40 ha are required to perform hunting using rifles. A shooting plan has to be submitted to the authorities and has to be approved by the Flemish Nature and Forestry Agency (ANB) although no minimum or maximum quota are in place for wild boar. Wild boar hunting (high seats or stalking) is allowed throughout the year and there are no limitations on sexes or ages. Flanders is divided in 10 wild boar management zones. Objectives for wild boar management are set within each management zone after consultation with local hunters, farmers, nature
organizations and other involved stakeholders. The objectives range from “no presence allowed” to “zero acceptance of damage” or “limited damage” as well as “local populations allowed but no further population increase”. For each wild boar shot, a hunting record has to be entered in the data portal of ANB. Hunters have to provide information on weight, sex and age class, in which game management unit the wild boar has been shot, etc. Moreover, hunters are obliged to preserve the lower jaw of shot animals. These are subsequently collected by the Institute for Nature and Forest Research (INBO) to sample DNA and determine age classes so population composition can be followed over the years.

**Distribution in Flanders**

Hunting bag numbers (reported number of hunted wild boar) are correlated to population numbers (assuming constant hunting pressure) (Casaer and Van Den Berge 2006, Massei et al. 2015, Myrberget 1988). Wild boar being a game species in Flanders, allows to follow the trend in population numbers and distribution based on reported hunting bags. With an increase from two wild boar hunted in 2006 to 1655 hunted in 2018, hunting bag and thus population numbers keep increasing exponentially.

Wild boar recolonization started in 2006 in the province of Limburg in and around the nature area ‘Zwarte beek’ in the Northwest and in and around the national park ‘Hoge Kempen’ in the East. Throughout the years, these two founder populations increased their distribution range resulting in presence throughout Limburg except from some southern municipalities. Distribution ranges also expanded towards the eastern municipalities of the provinces of Antwerp and Vlaams Brabant (Figure 1.2). Except for the settled population in and around Limburg, a few local sightings (and limited hunting bag) have occurred since 2006 in isolated locations outside the main distribution area. However, these sightings have not lead to established and growing populations.
Chapter 1

Figure 1.2: Distribution of wild boar with their current distribution range over municipalities in black (hunting bag statistics until 2018 in graph) and their distribution range over municipalities in the dashed area at the start of recolonization (hunting bag statistics 2007).

Crop damage

Due to intense interweaving of small natural areas, forest remnants, agricultural and urbanized areas in Flanders, wild boar come into close contact with human activities causing Human-Wildlife Impacts. Crop damage by wild boar draws high attention from farmers, hunters, nature managers and local authorities. The current extent of crop damage is unknown as there is no standardized monitoring method in place in Flanders and damage compensations for wild boar damage are only paid by the authorities in the exceptional case where damage is done by wild boar originating from a nature reserve or forest where no hunting or culling program is allowed. In all other cases the hunter himself is
responsible for paying for damage directly to the farmer. Although no data are available, both hunting and farming associations report that up till today hardly any damage was compensated by payments. However, sometimes hunters pay for electric fencing to prevent damage and/or place electric fences themselves.

A variety of different crops are damaged by wild boar including grasslands, maize, cereals, potatoes etc. Different studies on agricultural damage by wild boar report variable proportions of crop species being damaged; however grasslands and maize fields are often reported as the most preferred crops (Herrero et al. 2006, Lombardini et al. 2016, Schley et al. 2008). The amount of damage in a particular field can vary depending on the group size of foraging wild boar (large maternal group versus solitary males) and the number of days wild boar return to the same field. Preventive measures can offer a solution to reduce damage extent. Crop damage intensity has been found to be correlated with wild boar hunting bag thus reducing populations is seen as a useful management action (Geisser and Reyer 2004, Schley et al. 2008). The role of supplementary feeding to lead wild boar away from the fields has been debated as both positive as well as negative effects (as supplementary feeding can also cause a population increase) have been found (Bieber and Ruf 2005, Calenge et al. 2004, Geisser and Reyer 2004). A third method, often mentioned as an effective preventive measure which can be implemented is (electric) fencing (Geisser and Reyer 2004, Schley et al. 2008). The implementation of (electric) fencing as a preventive measure is also considered one of the conditions for receiving damage compensations in Flanders. However, the cost of fencing a field can be high limiting the implementation of such measures when farmers do not know the damage risk of a particular field.
Research questions

This PhD thesis aims to address knowledge gaps and to provide a sound scientific starting point in order to be able to conduct a risk assessment of wild boar impacts related to their potential future expansion in Flanders (Figure 1.3). Although wild boar is known to have a behavioural plasticity allowing them to adjust to anthropogenic landscapes, the Flemish context is quite different from most wild boar regions: Flanders has a human-dominated landscape and is characterised by an intense interweavement of small natural areas, forest remnants and agricultural areas, interspersed by a dense road network resulting in a severe level of landscape fragmentation. This leads to the question if wild boar expansion and foraging patterns are affected by this high degree of anthropogenic pressure on the landscape. If there might be limits towards the behavioural plasticity of wild boar, this can affect distribution mechanisms as well as crop damage patterns due to the inhibiting effect of fragmentation or by avoiding human presence. This induces uncertainty when research results from less anthropogenic areas on distribution or crop damage mechanisms would be projected to the Flemish situation. Research towards wild boar impacts and their expansion mechanisms in this type of landscape is essential to minimise this uncertainty. By getting better knowledge on current and potential crop damage distribution and the effect of fragmentation on damage risk, the extent of crop damage as a Human-Wildlife Impact will become clear. Moreover, we aim to use the gained information to assess how management tools can be implemented more efficiently to prevent these Human-Wildlife Impacts. Connecting this information with better knowledge on landscape connectivity and remaining habitat suitability in a highly fragmented landscape, will give an insight in future impacts with potential further expansion of wild boar in Flanders. Findings of this research can be applied to similar situations outside of Flanders as this research aims to provide knowledge on the extent of wild boars’ behavioural flexibility in fragmented landscapes.
As information on where crop damage occurs in Flanders was unknown at the start of this PhD, we started the PhD research by mapping crop damage distribution. In order to get these results quickly, we used the results of an email survey conducted by the University of Hasselt. This email survey was sent to farmers in the province of Limburg, given that this was the known distribution area of wild boar in Flanders at that time. We extended this survey with a non-respondent analysis, using telephone calls, as described in Chapter 2. The information gained in this first step was subsequently used to select a study area based on spatial distribution of crop damage which formed the basis for Chapters 3 and 4. Crop damage is currently not monitored in Flanders and there is a growing need for a time-efficient, standardized and accurate method to assess crop damage. Therefore Chapter 3 focussed on the development of a method allowing to assess damaged fields meeting these required conditions.
Using a drone combined with machine learning modelling, we developed a highly accurate method which can now be applied for assessing damaged fields. Given that taking preventive measures to avoid crop damage to a particular field can involve high costs, there is a need to be able to predict which fields are more susceptible to damage than others. This formed the focus of Chapter 4 in which the dataset of damaged fields used in previous chapter were analysed. Here, we focussed on the combination of both landscape- and field specific factors (like fertilization, crop type selection, etc.) influencing damage probability of a field to get a better knowledge of drivers of crop damage by wild boar. This assessment can now allow farmers to adjust agricultural practices to mitigate damage risk and taking preventive measures more efficiently by targeting high risk fields.

To gain a better knowledge of the behavioural plasticity of wild boar dealing with habitat fragmentation in human dominated landscapes, we analysed the effect of fragmentation and landscape connectivity, affecting future expansion of wild boar in Flanders in Chapter 5. For this purpose we conducted a landscape genetic analysis based on DNA-samples of the Flemish hunting bag. As these samples have been collected since the beginning of the recolonization of wild boar in Flanders, the influence of land cover types and habitat fragmentation on the recolonization process could be analysed using genetic structure analyses tools. Next, in Chapter 6, future distribution of wild boar was assessed by defining remaining habitat suitability outside their current distribution range. Occurrence data originated from citizen scientists such as waarnemingen.be, an observation platform for species records, as well as from hunting bag. These data allowed us to get a detailed insight in land cover types influencing habitat preferences of wild boar in Flanders. Finally, in Chapter 7 the general discussion covering and linking the results from the different chapters is provided. Connecting the results from all chapters’ results in a first risk assessment for both the current and future distribution of wild boar in Flanders (Figure 1.3). Conducting such risk assessment allows us to assess current and potential impact of wild boar guiding effective management planning and decision taking for management actions.
Chapter 2

Using an online survey to assess the spatial distribution of wild boar crop damage and factors influencing this distribution

Rutten, Anneleen1,2; Casaer, Jim2; Onkelinx Thierry2; De Smet, Lieven2; Witters, Nele3; Huysentruyt, Frank2; Leirs, Herwig1

1Evolutionary Ecology Group, Department of Biology, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium

2Research Institute for Nature and Forest (INBO), Havenlaan 88 bus 73, B-1000 Brussels, Belgium

3Environmental Economics Group, Centre for Environmental Sciences (CMK), Hasselt University, Martlarenlaan, B-3500 Hasselt, Belgium

DOI: 10.26496/bjz.2019.26
Abstract

Wild boar (Sus scrofa L.) reappeared in Flanders (northern Belgium) in 2006 after almost half a century of absence. Interactions between wild boar and human activities are frequent due to extensive fragmentation of the landscape in Flanders. Complaints about agricultural damage are increasing but the actual extent of crop damage remains unknown. The goal of this study was to assess the current risk and the spatial distribution of crop damage, as well as factors influencing damage distribution in the province of Limburg (eastern Flanders). An online survey was sent to farmers by email. Moreover, as we expected potential respondent bias towards farmers that already experienced damage, we also conducted a follow-up non-respondent check by telephone. Our study showed that the current crop damage probability on a farm lies between 42% (likely an overestimation due to respondent bias in the online survey) and 22% (an underestimation based on the non-respondent check). There is considerable geographical variation in the proportion of farms that report boar damage; probability for crop damage due to wild boar is relatively high for farmers in Limburg but shows a geographically heterogeneous spread. Factors explaining the crop damage probability differed strongly between the online survey and the non-respondent check and no consistent results could be found. Our results show that using the online survey, it was possible to get an initial insight in the geographical distribution of crop damage. However, as we found differences between the results of the online survey and the non-respondent check, taking management decisions based solely on online survey results without conducting a non-respondent check could lead to misguided actions.

Key words: Human-Wildlife Impact, online survey, damage probability, respondent bias
Introduction

The success of many wildlife populations depends on their capacity to coexist with humans (Carter and Linell 2016). Resolving Human-Wildlife Impacts (HWI, here defined as impacts due to interactions between wildlife and humans activities according to (Redpath et al. 2013)) is therefore seen as a major challenge for conservation (Treves et al. 2006). Species that re-establish in their former distribution area can typically be a source of conflict due to human safety risks or damage to crops, livestock, buildings, infrastructure, gardens, cars, etc. (Trouwborst 2010). One such re-established species causing impacts in Europe is the wild boar, *Sus scrofa* (Linnaeus, 1758, Mammalia, Chordata). Wild boar populations have been increasing all over Europe since the 1960s (Acevedo et al. 2007, Bieber and Ruf 2005, Saez-Royuela and Telleria 1986) and their expansion is still on-going in most European countries (Massei et al. 2015). A combination of environmental and agricultural changes have triggered the revival of the wild boar, mainly due to 1- increased piglet survival due to rising winter temperatures and 2- a lower age of first reproduction and increased average litter size related to increasing food availability. The latter factor is caused by more frequent mast years, agricultural intensification, decrease of grasslands and an increase in rapeseed and maize cultivation (Massei et al. 2015, Morelle et al. 2016).

The return of wild boar to Flanders (the northern part of Belgium) since 2006, after more than half a century of absence, rapidly developed from a few sightings to a now settled population. This is in line with the observed expansion in the rest of Europe (Scheppers et al. 2014) and is of interest because of the intense intertwinement of urbanized, agricultural and natural areas leading to close contact of wildlife with human activities. Flanders is currently one of the most densely populated areas in Europe (462 persons per km²) (Linell et al. 2001). Forests represent a mere 11.4% of the total surface area (Casaer and Licooppe 2010, Demolder et al. 2014) and the dense road network (5.2 km/km²) results in highly fragmented natural areas.
The key to successful wildlife management is to include both ecological and human dimensions (Keuling et al. 2016) into the decision making process in order to achieve positive impacts that are valued by all stakeholders (Riley et al. 2003). This is particularly true for the management of conflict species such as wild boar. Wild boar impacts can result in large costs and can be unacceptable for multiple stakeholders, with crop damage (including rooting in grasslands) as the primary impact (Cahill et al. 2012). In the Netherlands, France, Luxembourg and Wallonia (southern Belgium), the economic impact of crop damage by wild boar can be large (Carnis and Facchini 2012, Faunafonds 2014, Schley et al. 2008, Widar and Luxen 2016). In Flanders, compensation for wild boar damage is only paid by the authorities in the exceptional case where damage is done by wild boar originating from a nature reserve or forest where there is no hunting or culling program. In all other cases the hunter himself is responsible for paying for damage directly to the farmer. One of the consequences is that currently there is no centralised database covering damage paid annually. During annual meetings of wild boar management units, the need for better knowledge on the current extent of wild boar damage is stressed consistently by both farmers and hunters as well as nature administrators. Consequently there is an urgent need for a first assessment of current wild boar damage risk, its geographic distribution and factors driving it.

In this study, we aimed to assess:

- The probability of a farmer suffering wild boar damage given the current distribution of wild boar in Flanders.

- The spatial distribution of damage probability and factors influencing this distribution.

For this purpose, an online survey was implemented. Online surveys related to wildlife issues are used extensively as they can be easily implemented and they provide almost immediate results (Frederik 1998, Lard et al. 2002, Moore et al. 1999, Reiter et al. 1999, Signorille and Evans 2007, Stone 1973, White et al. 2005) Online surveys are also an inexpensive data-collection method (Duda and...
Nobile 2010, Sexton et al. 2011) and sending an online survey by email allows for flexibility as respondents can participate when it is convenient (Evans and Mathur 2005). Moreover, online surveys allow for direct data transfer into databases, eliminating manual input from questionnaires (Evans and Mathur 2005, Lefever et al. 2007). However, online surveys also present some drawbacks that call for caution. Sample validity, respondent and stakeholder bias and unverified respondents could all lead to inaccurate, unreliable and biased results (Duda and Nobile 2010, Evans and Mathur 2005, Lefever et al. 2007, White et al. 2005). Respondent bias occurs when the population of non-respondents differs from the respondent population. In this study respondent bias could have important consequences if farmers who had already suffered crop damage were more keen to respond than those that did not. Therefore we also conducted a non-respondent check.

Methods

Study area

At the onset of this study in 2015, the distribution of wild boar in Flanders was mostly limited to the province of Limburg (Figure 2.1) i.e. the eastern part of Flanders near the Dutch border (Scheppers et al. 2014). The geographical scope of this study was therefore limited to Limburg. With 23.6% of its surface area covered by forest, Limburg is the most forested province in Flanders (Demolder et al. 2014, Moors 2014), although still highly fragmented. Arable land in Limburg is dominated by maize (44.7%), grassland (32.3%) and cereal crops (23.0%) (Flanders: 40.8% - 38.0% - 21.2%) (Report Federale Overheidsdienst Economie (FOD), 2015).
Data collection

Online survey

The online survey was conducted in collaboration with Boerenbond, the largest cooperative farmers’ organization in Flanders, who provided farmer contact details. Of the 2824 farmers in Limburg (data STATBEL), 1282 are members of Boerenbond, 904 of whom had a known email address. Since we expected no correlation between damage probability, Boerenbond membership and whether or not an email address was known, we assumed the sampling population as representative for the total Limburg farmer population. In order to avoid an unverified responder bias, typically linked to an open access online survey, we sent farmers a personal invitation using their email-address to participate. All farmers with a known email address received the online survey invitation in March 2015. Reminders were sent after one and after two weeks to increase the response rate. The survey consisted of eleven questions (Supplementary materials, Appendix 2.1) and was estimated to take about 10 minutes to
complete. Only the answers to three questions (damage/no damage, municipality and farm size) were used for this research.

Non-Respondent check

In our study we acknowledged the possibility that farmers who had suffered crop damage would be more keen to respond than would those who had not. Moreover, a non-respondent check should always be conducted in cases where low response rates are expected (Duda and Nobile 2010, Sexton et al. 2011). Therefore, for each of the 43 municipalities, we randomly selected three farmers among the non-respondents from the online survey, contacted them by telephone and asked two questions from the questionnaire used for this study (damage/no damage and farm size).

Statistical analysis

All analyses were performed using R software (R Core Team 2016).

Survey sample

A series of Kendal rank correlation tests were conducted. Firstly to control whether the distribution of the farmers who received the online survey matched the distribution over the different municipalities of the target population (all farmers of Limburg). Next, we tested the spatial representativeness of our respondents to the online survey by comparing the distribution over municipalities of respondents and the distribution of the entire Limburg farmer population. Also the distribution of farm sizes of the respondents was compared to the total distribution of farm sizes in the province of Limburg.

Damage probability and its spatial distribution

Damage probability was defined as the proportion of farms that had experienced damage over the total number of farms per municipality and these values were displayed on a map to visualise spatial patterns. We conducted a Chi-square test to see whether damage probability differed between those who answered the online survey and the participants of the non-respondent check.
Factors influencing crop damage probability

The following analyses were done both for the online survey and for the non-respondent check. Doing this parallel analysis, we were able to assess a potential respondent bias and its impact on our findings.

To explore the factors influencing crop damage probability, apart from farm size (which was included as an offset factor to account for size effects) and the random effect of municipality, four additional potential explanatory factors were selected (Table 2.1).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crop damage</td>
<td>CD</td>
<td>Did farmers experience damage (Yes-No)</td>
</tr>
<tr>
<td>Farm size</td>
<td>FS</td>
<td>Three categories such as used in the survey: 0-40 ha, 40-60 ha, &gt;60ha</td>
</tr>
<tr>
<td>Municipality</td>
<td>MUN</td>
<td>Municipality in which most of the cropland of the farm is situated</td>
</tr>
<tr>
<td>Forest cover</td>
<td>FC</td>
<td>Percentage of forest cover per municipality (mean = 21.1%, SE = 13.1%)</td>
</tr>
<tr>
<td>Agriculture cover</td>
<td>AC</td>
<td>Percentage of agriculture cover per municipality (mean = 43.5, SE = 20.8)</td>
</tr>
<tr>
<td>Years of hunting of wild boar</td>
<td>YH</td>
<td>Number of years in which wild boar were shot in a municipality over the period 2007 – 2014 (mean = 2.74, SE = 0.15)</td>
</tr>
<tr>
<td>Number of wild boar shot</td>
<td>NS</td>
<td>Log (x+1) of number of wild boar shot over the period 2007-2014 (mean= 11.9, SE= 1.9)</td>
</tr>
</tbody>
</table>

Forest cover was included as the wild boar is a forest-dwelling species (Briedermann 1990). Previous research has also shown that forest cover influences wild boar distribution (Morelle et al. 2016) and serves as a good predictor of crop damage (Ficetola et al. 2014, Lombardini et al. 2016). The percentage of arable land (agriculture cover) was included as it has been shown to be the most important factor next to population density and forest edge length in modelling crop damage probability (Bleier et al. 2012). There are various suggested methods to estimate population densities of wild boar.
(Acevedo et al. 2007, Focardi et al. 2001, Focardi et al. 2002, Massei et al. 1997b), however finding a reliable method to estimate population densities at the municipality scale remains difficult (Engeman et al. 2013, Scheppers and Casaer 2012). We therefore included both the number of years in which wild boar were shot in a given municipality over the period 2007 – 2014 (Scheppers et al. 2014), and the number of wild boar shot in that period, as proxies for the abundance of wild boar. We consider these proxies as reliable given that no temporal or spatial variation of culling or hunting effort was expected, no quotas for wild boar hunting exist, no large areas are exempted from hunting, and hunting takes place year-round.

We screened these four potential factors for collinearity combining Variance Inflation Factors (VIF) (Fox and Monette 1992) in which we allowed a maximum VIF value of 2 and a correlation of which we set the maximum at 0.7. This revealed that forest and agriculture cover were strongly negatively correlated. We opted to select forest cover in further analysis following the findings of Ficetola et al. (2014) and Lombardini et al. (2016). The number of years in which wild boar were shot and the number of animals shot were also highly correlated. We opted to include the number of wild boar shot as a factor in the model based on the fact that the number of years wild boar were shot showed the highest VIF (VIF = 5.19). The remaining factors forest coverage and number of wild boar shot had VIF values below 2 and a correlation below 0.7.

We defined four linear mixed effect models for crop damage probability (Table 2.2). All models included farm size as an offset factor and the random effect of municipality. Model 1 was the null model in which no additional factors were included. In models 2 and 3 we additionally included forest cover or the number of wild boar shot respectively. Finally, in model 4, the full model, both forest cover and the number of wild boar shot were included.
Table 2.2: Four a priori models. All models include Farm Size (FS) as an offset variable and the random effect of municipality (fMUN). Covariates Forest Cover (FC) and number of wild boar shot (NS) are also included in hypothesized model 2, 3 and 4.

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Model structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>CD ~ FS + fMUN</td>
</tr>
<tr>
<td>Model 2</td>
<td>CD ~ FS + FC + fMUN</td>
</tr>
<tr>
<td>Model 3</td>
<td>CD ~ FS + NS + fMUN</td>
</tr>
<tr>
<td>Model 4</td>
<td>CD ~ FS + FC + NS + fMUN</td>
</tr>
</tbody>
</table>

AIC values (Akaike Information Criterion) (Hu 2007) were used for model selection: in cases where the AIC values between two models differed by more than 2, the model with the lowest AIC values was selected as the best model. Effect size (mean divided by standard error) was calculated for each covariate in the final model. When the online survey and the non-respondent check would have comparable results, the AIC values should return the same a priori model, and the influence of covariates should be comparable.

Results

Survey sample

Of the 904 farmers contacted, 182 (20%) completed the online survey and an additional 122 farmers were contacted by telephone for the non-respondent check.

The correlation coefficient between the spatial distribution of the total population of farmers over the different municipalities in Limburg and the one of the members of the Boerenbond who received the survey was 0.95 (p<0.001). The correlation coefficient between the total population of farmers and farmers who responded to the online survey was 0.65 (p<0.001). The distribution of farm sizes among the online survey respondents was not
correlated with that of the total population of farmers (-0.84, p>0.05) as survey respondents included mostly smaller farm sizes.

**Damage probability and its spatial distribution**

Forty two% (76 out of 182) of the respondents to the online survey reported damage, whereas only 22% (27 out of 122) of the non-respondents reported damage. Damage probability differed significantly between the two groups of respondents (online survey and non-respondent check) (chi-square test, p<0.001). Therefore, spatial distribution maps were made separately for the online survey and the non-respondent check (Figure 2.2). The highest reported damage prevalence was found in the North of the province of Limburg, according to both the online survey and the non-respondent check. Southern municipalities showed low to no damage probability.

Figure 2.2: Spatial distribution of reported damage probability. Left: overview of number of farmers who responded to the online survey/who received the online survey; grey scale gradient represents percentage of reported damage and right: percentage of damage reported in the non-respondent check with three contacted farmers per municipality.
Factors influencing crop damage probability

The model selection results are shown in Table 2.3. For the online survey, model 3 (including number of boars shot) was the most informative (lowest AIC value); for the non-respondent check, this was the case for model 2 (including forest cover). However, in both cases, the difference in AIC-values with model 4 (combining farm size with both hunting and forest cover) was very small (<2) so the quality of that model is not significantly lower but model 4 showed forest cover to be non-significant for the online survey, and number of boars shot to be non-significant for the non-respondent check. Therefore, model 3 was selected as the best model for the online survey and model 2 for the non-respondent check.

Table 2.3: Model selection. AIC (Akaike Information Criterion) and p-values of Log likelihood Test for all four a priori hypotheses. * is the selected model.

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Online survey</th>
<th>Non-respondent check</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>p-value</td>
</tr>
<tr>
<td>Model 1</td>
<td>201.13</td>
<td>115.77</td>
</tr>
<tr>
<td>Model 2</td>
<td>194.56</td>
<td>0.003</td>
</tr>
<tr>
<td>Model 3</td>
<td>180.99*</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Model 4</td>
<td>182.94</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Damage probability increases as farm size increases for both the model of the online survey and the non-respondent check, confirming our size effect expectation (Table 2.4). For the online survey the number of boars shot was positively correlated with damage probability. The model for the non-respondent check showed increasing damage probability with increasing forest cover.
Table 2.4: Model parameters of model 3 for the online survey and model 2 for the non-respondent check: Mean, standard error (SE), confidence intervals and effect size of the fixed effects describing the relationship between crop damage, farm size (FS), forest cover (FC) or number of wild boar shot (NS). * are significant parameters, p-value < 0.05.

<table>
<thead>
<tr>
<th></th>
<th>Online survey</th>
<th>Non-respondent check</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SE</td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.25</td>
<td>0.76</td>
</tr>
<tr>
<td>Farm Size 40-60 ha</td>
<td>1.39</td>
<td>0.5</td>
</tr>
<tr>
<td>Farm Size &gt;60 ha</td>
<td>1.44</td>
<td>0.55</td>
</tr>
<tr>
<td>Forest Coverage</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Number wild boar shot</td>
<td>1.16</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Predictions based on the final models (for intermediate farm size (40-60 ha), Figure 2.3) show that according to the model from the online survey, the predicted crop damage probability increases considerably with a higher number of boars shot. According to the model from the non-respondent check, the damage probability curve increases with forest cover.
Figure 2.3: Prediction for damage probability based on final model of the online survey (left) and of the non-respondent check (right) for intermediate farm size (40-60 ha).

Discussion

As the socio-cultural context of wildlife management has changed in recent decades (Riley et al. 2003), successful wildlife management more than ever implies including both ecological as well as human dimensions (Keuling et al. 2016). With wild boar causing more and more Human-Wildlife Impacts in Flanders, there was a need for a fast and simple method to collect information on the most important impact: crop damage. By using an online survey we aimed to gain an initial overview of the actual damage levels and their spatial distribution in the province of Limburg, Flanders.

Damage probability according to the online survey was 42%, whilst the non-respondent check returned a lower damage probability of 22%. We assume actual damage probability to be situated between these two values. As we consider farmers who already suffered crop damage to be more keen to respond than those that did not- introducing a respondent bias- we believe that the online survey resulted in an overestimation. However, we also consider the non-respondent check to be an underestimation as the surveyed population of this non-respondent check is based on farmers who did not respond to the
online survey and consequently excludes those farmers who already experienced and reported crop damage using the online survey.

We consider the correlation between the spatial distribution of the online survey respondents and of the total farmer population high enough to be spatially representative of the population of interest. Both the online survey and the non-respondent check show a clear spatial differentiation of the crop damage risk within the province with northern municipalities showing higher damage probabilities compared to southern municipalities. These findings illustrate the usefulness of an online survey -although generally overestimating damage probability- to get a quick scan of the actual spatial distribution of damage in a region when no other information sources are available. These results on spatial distribution of crop damage were therefore used in the research of Rutten et al. (2018) to delineate a study area in which damage cases were photographed to develop a drone based method to objectively assess damaged area in a field.

The online survey and the non-respondent check models resulted in divergent conclusions with respect to factors explaining the crop damage probability. According to the online survey the number of wild boar shot can best explain damage probability while forest cover was not included in the final model. However, the model of the non-respondent check showed a clear positive effect of forest cover on damage probability and the effect of the number of wild boar shot was not retained in the final model. These opposite findings made us conclude that our approach does not allow any reliable conclusions to be drawn for this question.

Possible explanations for the opposite findings of the models may be our small sample size on the one hand and the correlation between forest coverage and the number of wild boar shot on the other hand; or a combination of both elements. The relation between increasing wild boar numbers with increasing forest coverage has been found in previous studies (Bieber and Ruf 2005, Melis et al. 2006). In our study, the correlation between forest coverage and number of wild boar shot was 0.68. Given that this value is smaller than the set threshold value of 0.7 used for eliminating variables from the model building, both variables
were included as factors in the model selection. Regarding the small sample size, we needed email addresses and telephone numbers of farmers, thus the collaboration with farmers’ organization Boerenbond was essential. This resulted in a total of 904 email-addresses, equalling 32% of the total farmer population in Limburg. However, we only achieved a response rate of 20% for the online survey.

While the goal of surveys like ours is to collect information beneficial for efficient management of Human-Wildlife Impacts (Browne-Nuñez and Jonker 2008), authors too often ignore the problem of respondent bias. According to a review on surveys in ecology by White et al. (2005) only 12.5% of surveys included any form of non-respondent check. Studies like that of Barker (1991) who did conduct such a check on the scope of hunter harvest of waterfowl in New Zealand, found that harvest estimates were about 20% higher according to respondents when compared to non-respondents; obviously, this can have large implications when bias is not accounted for when setting up survey-based harvest quotas. Potential consequences of not validating respondent bias also became clear in the present study: not only did factors influencing the spatial distribution differ between the online survey and the non-respondent check models, but we also found an overall higher damage probability according to predictions of the online survey model compared to the non-respondent check. This can lead to misguided conclusions and therefore misinformed decisions on management actions should they be based on these results only.

Conclusion

Using the online survey followed by a non-respondent check, we found clear spatial differentiation of crop damage risk within the province of Limburg; the online survey and the non-respondent check concurred in the geographic distribution of this risk. However we could not determine unambiguously the factors influencing damage probabilities as we found opposite results between the online survey and the non-respondent check. For future studies we
recommend caution with the use of online surveys, certainly in cases where the risk for a low response rate is high. We suggest a focus on increasing sample size as much as possible. This can be achieved by providing rewards for completing the survey, persuading respondents that their responses will be used, providing frequent reminders, keeping the survey short or enhancing the survey with visual elements. (Deutskens et al. 2004, Nulty 2008). However, we want to stress that at all times a non-respondent check should be conducted to assess the validity of online survey results.

Acknowledgements

We would like to thank Boerenbond for its cooperation and for providing contact details on farmers in the region. We also thank the Flemish Land Agency (VLM) for the data on the distribution of farms across municipalities in Limburg. We are very grateful to all the farmers who participated in the surveys. We are grateful to Kris Somers and Sebastien Lizin from Hasselt University for the setup of the online survey, data collection and preliminary analysis. Finally, we want to express our appreciation to dr. Dirk Maes, dr. Lucida Kirkpatrick, Bram Vanden Broecke, Frederik Van de Perre and Emma Cartuyvels for their valuable advice on the manuscript.
Chapter 3

Assessing Agricultural Damage by Wild Boar Using Drones

Rutten, Anneleen¹,²; Casaer, Jim²; Vogels, Marjolein A.³; Addink, Elisabeth A.³; Vanden Borre³, Jeroen; Leirs, Herwig¹

¹Evolutionary Ecology Group, Department of Biology, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium

²Research Group Wildlife Management, Research Institute for Nature and Forest (INBO), Havenlaan 88 bus 73, B-1000 Brussels, Belgium

³Faculty of Geosciences, Department of Physical Geography, Utrecht University, Princetonlaan 8a, NL-3584 CB Utrecht, The Netherlands

Abstract

In Flanders (northern Belgium), wild boar (*Sus scrofa*) returned in 2006 after 50 years of absence and the population is increasing, both in abundance and geographic extent. In the absence of wild boar, Flanders’ landscape structure changed into a dense, mosaic-like pattern of agricultural, natural, and urban areas. The return of the wild boar increasingly leads to Human-Wildlife Impacts, mainly linked to damage in agriculture. Hence, there is a growing need for a time-efficient, standardized, and accurate method to assess crop damage. We present an Unmanned Aerial Vehicle-based method, using Geographic Object-Based Image Analysis and Random Forests to estimate the damaged area and associated yield losses, between 2015 and 2017, due to wild boar in individual fields in Flanders. Our approach resulted in an 84.50% overall accuracy in calculating damaged area for maize fields and 94.40% for grasslands. Damage levels ranged between 14.3% and 20.2% in maize fields and 16.5% to 25.4% in grasslands. Our method can provide objective base data for compensation schemes and guide management strategies based on damage assessments.

**Key words:** Belgium, crop damage, GEOBIA, Geographic Object-Based Image Analysis, UAV, wildlife damage.
Introduction

The wild boar (*Sus scrofa*) is one of the most widespread mammal species of the world (Keuling et al. 2018, Massei and Genov 2004) with populations expanding throughout Europe since the 1960s (Acevedo et al. 2007, Bieber and Ruf 2005, Massei et al. 2015, Saez-Royuela and Telleria 1986). Expanding abundance of wild boar is challenging for both conservation (Barrios-Garcia and Ballari 2012) and society because Human-Wildlife Impacts (HWI, here defined as impacts due to interactions between wildlife and humans activities according to (Redpath et al. 2013)) arise linked to damage to crops (including rooting of grasslands), traffic collisions, and disease transmission (Amici et al. 2012, Bieber and Ruf 2005, Massei et al. 2015, Morelle et al. 2013, Treves et al. 2006). In Flanders (northern Belgium), where the wild boar was absent for more than half a century, the species returned in 2006 and its population is increasing rapidly (Scheppers et al. 2014). During the past few decades, landscapes in Flanders have changed dramatically due to urbanization, agricultural intensification, and fragmentation. Flanders has become one of the most densely populated areas in Europe: 462 persons/km² (FOD Economie 2011- unpublished report, Linell et al. 2001). A fragmented structure with a mosaic-like pattern composed of small natural areas, forest remnants, agricultural areas, and urbanized areas, which are all crossed by a dense road network (5.2 km/km²), characterizes the landscape. This results in a situation with frequent wildlife-human interactions and wildlife-related effects that warrant management attention (Riley et al. 2003).

Like many countries and regions in Europe, crop damage by wild boar is not monitored in Flanders and no compensations are paid. Therefore, we lack knowledge on the current extent of crop damage and associated losses for the agricultural sector. Yet, the magnitude of crop damage by wild boar can be significant as shown in some surrounding countries and regions (Carnis and Facchini 2012 for France, Faunafonds 2014 for the Netherlands, Schley et al. 2008 for Luxembourg, Widar and Luxen 2016 for Wallonia). However, assessments are done by a variety of methods because currently no well-
established and accepted method exists to assess damage in an accurate and objective manner (Michez et al. 2016). Moreover, farmers who are the most affected stakeholders report an increasing need for monitoring crop damage by wild boar. Consequently, there is a need for a standardized monitoring of crop damage by wild boar.

Any methods to assess crop damage should be standardized, objective, accurate, time-efficient, allow a full assessment, and applicable to different crops. Existing methods include ground visits with visual assessment, mapping damage spots with handheld Global Positioning System (GPS; Engeman et al. 2007a, Felix et al. 2014), and estimations extrapolated from randomly selected transects or plots (Chavarria et al. 2007, Cushman et al. 2004, Engeman et al. 2007b). These approaches do not meet all of the desired requirements because they are either time-consuming, subjective, or unsuitable for larger areas. A suitable method should be able to provide data both for damage compensation schemes and modelling socio-economic impacts of competing management strategies (Reyns et al. 2018).

The advantage of using photographs to assess rooting by wild boar was been shown by Engeman et al. (2016), who took photographs of damaged grasslands in the mountainous landscape of Romania from vantage points and assessed damaged area using Geographic Information System by manually outlining rooted areas. Although they showed this method to be quick and efficient, it can only be used in mountainous areas; but, they suggested that this method could also be applied using drones. The use of camera equipped Unmanned Aerial Vehicles (UAV), “drones” might indeed offer a practical solution. In recent years, use of drones has strongly increased because of easier access, flexible data acquisition possibilities and reduced costs (Salamí San Juan et al. 2014). Drones offer continuous coverage, collect data at centimetre resolution, require little training to operate, and can be deployed at short notice. Michez et al. (2016) and Kuželka and Surový (2018) recently showed how drones can be used to assess crop damage by wild boar in maize (corn) and wheat fields using generated photogrammetric digital elevation models from aerial photographs.
taken with a drone, where a threshold in height difference allowed them to
distinguish damaged from undamaged crops. However, Michez et al. (2016)
also outlined that this method is less applicable to crop types where damage
does not involve height difference like grasslands. A manual delineation would
be more applicable in croplands, but this is not objective and involves a time-
consuming procedure. Therefore, an automated processing flow is desired
(Engeman et al. 2016). Geographic Object-Based Image Analysis (GEOBIA),
in which pixels are grouped into informative objects, that is, coherent landscape
elements (Addink et al. 2012, Blaschke 2010), is a technique that can be used as
a standardized semi-automated method for interpretation of aerial photographs
Object-Based Image Analysis has been shown to be useful in assessing the
severity of crop damage by insects on sorghum crops (Puig et al. 2015) and
mapping cane grub (Dermolepida albohirtum) damage on sugarcane plants
(Johansen et al. 2014, Johansen et al. 2017). We investigated whether GEOBIA
can be an appropriate technique to analyse aerial photos of fields damaged by
wild boar in an accurate and semi-automated workflow.

We developed a semi-automated workflow to assess crop damage at the field
level on UAV imagery. Our objectives were 1) assessing the accuracy with
which crop damage can be calculated using GEOBIA; 2) assessing the variation
of damaged area in damaged fields; and 3) assessing the time and cost-
efficiency of damage estimation from UAV images.

Methods

Study area

Flanders (northern Belgium) had a highly fragmented land cover with 11.4%
forest coverage and 53% agricultural coverage (Demolder et al. 2014). The
distribution of wild boar in Flanders was largely limited to the eastern province
of Limburg and some eastern municipalities in the province of Antwerp
(Scheppers et al. 2014). In Limburg, the area where most damages were
reported by farmers was selected based on the results of an online survey (Rutten et al., unpublished data; Figure 3.1). Farmers within the study area could report crop damage by wild boar in the scope of this research. All reported damage cases from 2015, 2016, and 2017 were assessed and included in this study.

![Figure 3.1: Location of study area (blue) in the province of Limburg (dark grey) and Flanders (northern part of Belgium, light grey) in which 133 crop fields damaged by wild boar have been photographed via drone between 2015 and 2017. Dashed areas: distribution area of wild boar in Flanders in 2014.](image)

**Data Acquisition**

When farmers reported crop damage by wild boar, we photographed the damaged field using a UAV (DJI Phantom 3 Advanced using default included camera: 12 megapixel, f/2.8, 948 field of view; DJI, Shenzhen, China) just
before harvesting maize fields or shortly after reporting damage in grasslands. Using the Pix4D Capture App (Pix4D S.A., Lausanne, Switzerland) as a flight planner while flying at 40–45-m height, we took serial photos with 80–85% overlap between photos. Afterward we stitched photographs into a georeferenced orthophoto using Agisoft Photoscan (Agisoft LLC, St. Petersburg, Russia) or ENVI Onebutton (Icaros, Fairfax, VA). We clipped individual fields from the orthophotos to exclude the surrounding landscape from the analysis. In total, we photographed 133 damaged fields (Supplementary materials, Appendix 3.1).

**Geographic Object-Based Image Analysis**

*Principle*

Geographic Object-Based Image Analysis is a technique in which classification of the photographs is not based on pixels, but on objects. These represent groups of neighbouring pixels that are spectrally similar (Addink et al. 2010, Blaschke 2010). Subsequent classification of objects is not limited to spectral information (as is characteristic for pixel-based approaches), but classification can be based on information on overall colour and tone, texture, pattern, shape, shadow, context, and size of objects (Blaschke et al. 2014). This makes the workflow of GEOBIA similar to our human visual perception of the world, so coherent landscape elements can be defined and used for landscape classification (Addink et al. 2012). Our specific goal was to classify damaged fields into undamaged and damaged areas.

*Image segmentation*

Geographic Object-Based Image Analysis starts with a segmentation step in which we segmented orthophotos into objects representing meaningful landscape elements. We performed this segmentation using the eCognition Developer software (Trimble Inc., Westminster, CO, USA). We grouped pixels into homogeneous objects using multiresolution segmentation, which is based on a heterogeneity threshold considering both spectral similarity and shape
characteristics. In subsequent steps, we merged small objects into larger objects until the heterogeneity threshold is reached (Benz et al. 2004). We visually optimized the threshold such that the objects were as large as possible while representing either damaged or undamaged crops, avoiding a combination of the 2 classes (Figure 3.2).

![Illustration of segmentation (using eCognition) of a crop field damaged by wild boar, photographed via drone between 2015 and 2017 in the study area in Flanders (northern part of Belgium). Segmentation of orthophotos of damaged fields is the first step in Geographic Object-Based Image Analysis (GEOBIA) in which pixels are grouped in homogeneous objects using multiresolution segmentation. Left: Maize field orthophoto derived from drone photographs (A) and with segments outlined derived from the eCognition software (B). Right: grassland orthophoto derived from drone photographs (C) and with segments outlined derived from the eCognition software (D).]

**Random Forest Models**

We performed classification of objects using the Random Forest (RF) algorithm (Breiman 2001) based on a set of 25 attributes describing shape (4), texture (8), and spectral properties (13) of the objects (Supplementary materials, Appendix 3.2). The Random Forest is a robust classifier that makes predictions based on a training set (independent variables) using multiple decision trees. Random Forests are increasingly used in land-use and land-cover classifications (Rodriguez-Galiano et al. 2012). We created the Random Forests in this study using the `randomForest` R package (Liaw and Wiener 2002) available in the R
software environment (R Studio, Boston, MA, USA). We built a RF-model separately for maize and grasslands, with the number of trees set to 10,000 for each model. We initially trained the RF-models on information from the maize fields and grasslands of 2016. Ideally, the RF-model would classify data from any year with similar accuracy values, without the need for calibration when a new set of photos arrives. When the accuracy for the 2016 RF-model was rather low (we set an arbitrary limit of 80% overall accuracy, which we regarded the absolute minimum), we added fields from other years (2015 and 2017) to include interannual variation because this notably influences model performance. For maize fields, preliminary model building indeed showed that based on 22 segmented orthophotos of maize fields of 2016, model performance did not reach 80% overall accuracy; therefore, we added 5 extra maize fields of 2015 and 5 of 2017 resulting in 32 total fields. This was not the case for grasslands because model performance reached the 80% threshold of accuracy, so we only used 26 fields of 2016 for the RF-G model.

For each field, we constructed a training and a validation data set by visually interpreting randomly selected objects: in each field we assigned ≥ 100 objects to damage, 100 to crop (maize or grass), 100 to bare soil (for maize fields only, to differentiate damage from bare soil between maize rows), and 100 objects to undamaged areas shaded by nearby trees (for grasslands only and only if shadow was present in the orthophoto; in these cases, we also included sufficient damaged objects in the shaded area to incorporate the difference between shaded and unshaded damage). In total, we assigned 5,292 objects to maize, 3,802 to damage, and 5,048 to soil in the 32 maize fields. In the 26 grasslands, we allocated 3,700 objects to grass, 3,126 to damage, and 373 to shadow. Subsequently, we used 70% of these objects for training the RF-model. We derived variable importance for each attribute, expressed by MDA (Mean Decrease in Accuracy, Supplementary materials, Appendix 3.2), which is a measure of loss of accuracy when the variable is left out (Cutler et al. 2007).
Validation Measures

To evaluate the accuracy of the model-deduced damage maps, we used 3 types of validation measures (Figure 3.3). We grouped the categories maize and soil (for maize fields), and grass and shadow (for grasslands), in the single category “no damage” because their individual accuracies were not of interest to the study. This allowed for a binary accuracy assessment of “damage” versus “no damage.”

Figure 3.3: Workflow used for the calculation of the accuracy measures of the damaged area, which is derived from Geographic Object-Based Image Analysis (GEOBIA) and Random Forest models, to estimate the damaged area in 133 crop fields damaged by wild boar, which were photographed via drone between 2015 and 2017 in Flanders (northern part of Belgium).

Validation of the model

We used the remaining 30% of the labelled objects (i.e., other than the 70% training set) for validation of each model. We calculated a confusion matrix with a set of accuracy measures corrected for object area using the binary assignment. The accuracy measures include user’s accuracy, which is the area of correctly classified objects of a class divided by the total area of predicted objects in a class; producer’s accuracy, which is the area of correctly classified objects of a class divided by the total area of reference objects in a class; and,
overall accuracy and the kappa coefficient, which is a measure of how well the model performed compared with performance by chance (Cutler et al. 2007).

**Performance as a crop-damage assessment tool**

We assessed the value of the model for its practical application to evaluate damage for newly collected field imagery by testing its performance on all fields not used for model construction. We manually assigned $\geq 50$ objects on each of these independent fields to 1 of 3 possible classes (i.e., damage, crop, soil in maize fields; and damage, crop, shadow in grasslands). We then used this set of objects to validate the RF-models. We set up a confusion matrix and we calculated the same accuracy measures as mentioned corrected for object area size.

**Ground-truthing**

We provided a third validation measure by a ground-truthing check. In 10 maize fields and 10 grasslands photographed in 2017, we took 10 GPS locations for damage and 10 GPS locations for undamaged crop (maize or grass) using a Trimble advanced RTK R6 (Trimble, Sunnyvale, CA, USA; 0.012-m horizontal accuracy on average) on the same day that we photographed the field. This resulted in 400 ground-truthing points. We overlaid these GPS locations with the corresponding object after segmentation and classification of the orthophoto by the RF-model. Based on this comparison, we set up a third confusion matrix and we calculated accuracy measures (not based on area calculations but on presence–absence of damage).

**Assessment of damaged area**

Using accuracy measures of corresponding model classes, we could calculate damaged area and damaged percentage of a field (Figure 3.4). The error on these calculations involved both user’s and producer’s accuracy of the objects classified as damage and was calculated in 2 steps. First, we calculated the true positive rate (TPR) of damaged area (percentage of damaged area that were correctly classified as damaged) by multiplying damaged area by the user’s accuracy (UA; $\text{TPR} = \text{area} \times \text{UA}$). Secondly, we corrected the damaged area for
the false negative rate (FNR) using the producer’s accuracy (PA): percentage of damage area which was probably missed (FNR=1/PA). The error is thus calculated using the following formula:

\[ \text{Error} = TPR \times FNR \]

*Figure 3.4: Workflow for calculating the error on damaged area, which is derived from Geographic Object-Based Image Analysis (GEOBIA) and Random Forest models, to estimate the damaged area in 133 crop fields damaged by wild boar, which were photographed via drone between 2015 and 2017 in Flanders (northern part of Belgium). The error on the damaged area is calculated using accuracy measures of corresponding classes (model validation or model performance), true positive rate (TPR), user’s accuracy (UA), and producers’ accuracy (PA).*
Results

Validation and Performance of the Model

Maize fields

Model validation (i.e., 30% of objects of the fields used for model construction) showed a high overall accuracy of 96.45%, whereas the model performance with an accuracy of 84.50% shows that classification is more difficult for fields not used for model construction (Table 3.1). However, ground-truthing showed 94.50% overall accuracy for the constructed RF-model.

Grasslands

In the final RF-model for grasslands, 26 orthophotos of 2016 were used (model performance exceeded an overall accuracy of 80% only using fields of 2016). For grasslands, model validation shows an overall accuracy of 95.71% and model performance resulted in an accuracy of 94.40% (Table 3.2). Ground-truthing showed 98.00% accuracy.
Table 3.1: Confusion matrices (a¼area [m²] and between brackets: n¼no. of objects) and accuracy measures (corrected for object area) for the Random Forest model for assessing maize fields damaged by wild boar, for which 79 damaged maize fields were photographed via drone between 2015 and 2017 in Flanders (northern part of Belgium), and a Geographic Object-Based Image Analysis and Random Forest model was developed.

<table>
<thead>
<tr>
<th>Model validation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>a=91.20m²</td>
<td>n=4,243 objects</td>
</tr>
<tr>
<td>Predicted</td>
<td></td>
</tr>
<tr>
<td>Damage</td>
<td>29.62 (1,065)</td>
</tr>
<tr>
<td>No damage</td>
<td>1.71 (76)</td>
</tr>
<tr>
<td>Producer’s accuracy (%)</td>
<td>94.55</td>
</tr>
<tr>
<td>Overall accuracy (%)</td>
<td>96.45</td>
</tr>
<tr>
<td>Kappa coeff.</td>
<td>0.92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model performance</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>a=39.12m²</td>
<td>n=2,422 objects</td>
</tr>
<tr>
<td>Predicted</td>
<td></td>
</tr>
<tr>
<td>Damage</td>
<td>21.52 (990)</td>
</tr>
<tr>
<td>No damage</td>
<td>4.26</td>
</tr>
<tr>
<td>Producer’s accuracy (%)</td>
<td>83.49</td>
</tr>
<tr>
<td>Overall accuracy (%)</td>
<td>84.50</td>
</tr>
<tr>
<td>Kappa coeff.</td>
<td>0.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ground-truth</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n=200 objects)</td>
<td></td>
</tr>
<tr>
<td>Predicted</td>
<td></td>
</tr>
<tr>
<td>Damage</td>
<td>95</td>
</tr>
<tr>
<td>No damage</td>
<td>5</td>
</tr>
<tr>
<td>Producer’s accuracy (%)</td>
<td>95.00</td>
</tr>
<tr>
<td>Overall accuracy (%)</td>
<td>94.50</td>
</tr>
<tr>
<td>Kappa coeff.</td>
<td>0.89</td>
</tr>
</tbody>
</table>
Table 3.2: Confusion matrices (a¼area [m$^2$] and between brackets: n¼no. of objects) and accuracy measures (corrected for object area) for the Random Forest model for assessing damaged grasslands by wild boar, for which 54 damaged grasslands were photographed via drone between 2015 and 2017 in Flanders (northern part of Belgium), and a Geographic Object-Based Image Analysis and Random Forest model was developed.

<table>
<thead>
<tr>
<th>Model validation a=92.54m² (n=2,160 objects)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted</td>
<td>Damage</td>
</tr>
<tr>
<td>Damage</td>
<td>21.45 (857)</td>
</tr>
<tr>
<td>No damage</td>
<td>2.12 (81)</td>
</tr>
<tr>
<td>Producer’s accuracy (%)</td>
<td>91.03</td>
</tr>
<tr>
<td>Overall accuracy (%)</td>
<td>95.71</td>
</tr>
<tr>
<td>Kappa coeff.</td>
<td>0.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model performance a=47.23m² (n=1,245 objects)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted</td>
<td>Damage</td>
</tr>
<tr>
<td>Damage</td>
<td>10.44 (515)</td>
</tr>
<tr>
<td>No damage</td>
<td>1.64 (75)</td>
</tr>
<tr>
<td>Producer’s accuracy (%)</td>
<td>86.39</td>
</tr>
<tr>
<td>Overall accuracy (%)</td>
<td>94.40</td>
</tr>
<tr>
<td>Kappa coeff.</td>
<td>0.85</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ground-truth (n=200 objects)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted</td>
<td>Damage</td>
</tr>
<tr>
<td>Damage</td>
<td>99</td>
</tr>
<tr>
<td>No damage</td>
<td>3</td>
</tr>
<tr>
<td>Producer’s accuracy (%)</td>
<td>97.01</td>
</tr>
<tr>
<td>Overall accuracy (%)</td>
<td>98.00</td>
</tr>
<tr>
<td>Kappa coeff.</td>
<td>0.96</td>
</tr>
</tbody>
</table>
**Crop Damage Assessment**

The average damaged area (see Figure 3.5 for examples) was 17.2% in maize fields and 20.6% in grasslands (Table 3.3). Using corresponding accuracy measures, the error on damaged area could be calculated as well as the error on the damaged percentage of a field (Supplementary materials, Appendix 3.3). In terms of time- and cost-efficiency of our drone method, we made a comparison (Table 3.4) with ground-based estimations as applied in Wallonia (southern Belgium; J. Widar, Fourrages Mieux, personal communication). Start-up costs for our presented drone method are lower than the method of ground-based estimations (Table 3.4).

*Table 3.3: Average percent damaged area within a field (with min. and max.) over all 133 fields damaged by wild boar, which were photographed via drone between 2015 and 2017 in Flanders (northern part of Belgium), as derived from a Geographic Object-Based Image Analysis and Random-Forest model classification that was developed.*

<table>
<thead>
<tr>
<th>Crop</th>
<th>Year</th>
<th>No. of fields</th>
<th>Average percent damaged (min.-max.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maize</td>
<td>2015</td>
<td>21</td>
<td>16.50 (0.83-44.89)</td>
</tr>
<tr>
<td></td>
<td>2016</td>
<td>26</td>
<td>12.70 (0.36-45.11)</td>
</tr>
<tr>
<td></td>
<td>2017</td>
<td>32</td>
<td>22.40 (2.09-52.87)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>79</td>
<td>17.20</td>
</tr>
<tr>
<td>Grass</td>
<td>2015</td>
<td>1</td>
<td>19.00</td>
</tr>
<tr>
<td></td>
<td>2016</td>
<td>33</td>
<td>19.10 (3.98-87.77)</td>
</tr>
<tr>
<td></td>
<td>2017</td>
<td>20</td>
<td>24.00 (4.11-48.82)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>54</td>
<td>20.60</td>
</tr>
</tbody>
</table>
Figure 3.5: Visualization of the resulting Geographic Object-Based Image Analysis (GEOBIA) Random-Forest model classification of the area damaged by wild boar (yellow) in a maize field (A=original field, B=classified damage) and grassland (C=original field, D=classified damage) as 1 of the 133 crop fields damaged by wild boar photographed via drone between 2015 and 2017 in Flanders (northern part of Belgium).

The labour time for a field visit and damage processing varies widely depending on the accuracy of ground-based estimation (from 90min for an estimation of 10% of a damaged field of 5 ha to nearly 26 hr for a full exhaustive assessment), whereas the labour time and accuracy is fixed using the drone method (150 min for the same field).
Table 3.4: Overview of costs and time requirement involved in damage assessment of fields damaged by wild boar during 2015 and 2017 using the presented drone method or ground visit like done in Wallonia (southern part of Belgium, J. Widar, Fourrages Mieux, personal communication), in which only a part of the fields is estimated depending on the intensity of the damage. Hourly wages are not included as these can be variable. Passive processing time is not included as this does not influence active labour time. € = euros.

<table>
<thead>
<tr>
<th>Drone</th>
<th>Start-up costs</th>
<th>Ground visit</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENVI Onebutton license</td>
<td>€890</td>
<td>Estimation software development: €25,000</td>
</tr>
<tr>
<td>Agisoft Photoscan Pro licence</td>
<td>€2,900</td>
<td></td>
</tr>
<tr>
<td>eCognition licence</td>
<td>€1,716</td>
<td></td>
</tr>
<tr>
<td>DJI Phantom 3 Advanced</td>
<td>€1,200</td>
<td></td>
</tr>
<tr>
<td>Sufficient batteries for a full day field assessments</td>
<td>€1,500</td>
<td></td>
</tr>
<tr>
<td>Total cost</td>
<td>€8206</td>
<td>Total cost</td>
</tr>
</tbody>
</table>

Field visit

<table>
<thead>
<tr>
<th>Photographing field</th>
<th>5 min/ha</th>
<th>Ground assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Surveying 10%:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15 min/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Surveying 20-25%:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30 min/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Full assessment:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 hr and 6 min/ha</td>
</tr>
</tbody>
</table>

Data processing

1) Stitch photos | ± 2 hr/field | Applying estimation software 15 min/field
2) Segment photos |          |                     |
3) Apply RF-model |          |                     |

Total field of 5 ha | 2 hr 30 min | Total field of 5 ha | From 1 hr 30 min to 25 hr 45 min

Discussion

Our presented method applies machine learning using GEOBIA on UAV imagery of damaged crop fields by wild boar to calculate damaged area, which is shown to be an objective, time-efficient, and accurate approach. Model performance showed high overall accuracies, with greater accuracies for grasslands than for maize fields. We consider the presented method to be useful as a tool to get a detailed and objective estimation of damaged areas in maize and grasslands.
To reach an acceptable model performance of >80% (as an arbitrary limit we set), we needed to combine maize field data from several years, indicating a larger variation among maize fields. Including fields from >3 years might improve model performances for damaged maize fields substantially. Given the high accuracy for grasslands, the expected gain when adding grasslands for >1 year seems not sufficiently advantageous compared with the required time investment.

We applied our method to maize and grasslands because we did not have sufficient reported damage cases for other crops. As long as damage is visually distinguishable in aerial photographs (Addink et al. 2012, Blaschke 2010), we are confident that our method can be applied to other crops such as wheat, oats, etc. Michez et al. (2016) pointed out that object-based image analysis improved their classification method in which they used digital elevation models and height thresholds to distinguish damaged from undamaged crops. We studied damage from wild boar (ground check in all cases), but other damage causes do exist, such as those caused by other wildlife species (e.g., badger Meles meles). These sources of damage can have similar visual characteristics in aerial photographs and might be distinguishable by our GEOBIA-RF model as well. We did not have any cases of damage by other species, so we could not check this nor the possibility to distinguish damage sources.

In the assessed maize fields we found that, on average, 17.2% of the area in fields was damaged, whereas in grasslands this figure was 20.6%, although a large variation was found for both crops. Bueno et al. (2010) reported 16% of the assessed area in damaged livestock pastures in Spain to be uprooted (plants pulled out of soil), Engeman et al. (2016) found between 11.2% and 13.5% of the damaged grasslands in Romanian mountains to be rooted. Bueno et al. (2009) reported up to 12% of the total areas of damaged Pyrenean alpine and subalpine grasslands to be actually damaged.

Using local crop prices (average yield (€/ha) over the period 2013–2017 in Flanders, prices according to local farmers’ union Boerenbond), direct yield losses in maize fields would be on average 342 €/ha of maize field and 282 €/ha
of Grasslands. However, this is only a rough estimate because regional and year-dependent yield differences are not taken into account. Moreover, economic losses to farmers are likely greater than merely yield losses. For example in grasslands, uneven surfaces as a result of wild boar rooting may cause damage to mowing machines and restoration measures are needed to repair the damaged grassland (Frederik 1998).

The economic effect of damage is the main limiting factors in stakeholders’ tolerance toward wildlife (Carpenter et al. 2013). Compensation schemes may increase tolerance of wildlife and promote more positive attitudes toward concerns and therefore decrease the number of Human-Wildlife Impacts (Nyhuis et al. 2005). However, often there is little quantitative evidence and costs are mostly estimated (Nyhuis et al. 2003, Nyhuis et al. 2005). When comparing labour time of ground-based estimations with our drone method, we see that time use depends on the accuracy of the ground-based estimation. Start-up costs seem accountable to us when compared with figures like our reported cost per hectare, given that our study area was >330 ha, and the reported damage costs in Europe alone is approximately 80 million euros annually (Putman and Apollonio 2014).

Acknowledgements

A. Rutten holds a PhD grant (IWT.141377) from the Agency Flanders Innovation & Entrepreneurship (VLAIO), Flanders, Belgium. We thank K. Mertens and P. Lootens from the Flemish Research Institute for Agriculture, Fisheries and Food (ILVO) and J. De Reu from the Flemish Research Institute for Nature and Forest (INBO) for their help with photo pre-processing, and Prof. Dr. Ir. B. Somers from KULeuven for methodological suggestions. Special gratitude goes to the Associate Editor and reviewers for their contributions.
Additional remarks in the dissertation after publication of the paper

It should be noted that table 3.4 about the costs and time requirement for the damage assessment using the drone method or ground visit can be misleading. While the time needed to develop the estimation software for the ground visit method is included in the development price, the time investment to develop the drone-based model as part of this PhD. was not included. Moreover, we are convinced that additional training will be needed for the drone method to account for variations throughout years, resulting in additional costs that are currently not included in this table.

One could wonder why we did not use Digital Elevation Models (DEM) that can be obtained using the Agisoft photoscan software. However, as crop damage by wild boar only resulted in a height difference in maize fields (due to broken stems) but not in grasslands (height difference from rooting is minimal), and our goal was to develop one robust method using drone photographs for damage assessments in all kinds of agricultural fields, we decided not to use DEM in our approach.
Chapter 4

What attracts wild boar to agricultural crops?

Rutten, Anneleen\textsuperscript{1,2}; Casaer, Jim\textsuperscript{2}; Strubbe, Diederik\textsuperscript{3}; Leirs, Herwig\textsuperscript{1}

\textsuperscript{1}Evolutionary Ecology Group, Department of Biology, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium

\textsuperscript{2}Research Institute for Nature and Forest (INBO), Havenlaan 88 bus 73, B-1000 Brussels, Belgium

\textsuperscript{3}Terrestrial Ecology Unit, Ghent University, K. L. Ledeganckstraat 35, B-9000 Gent, Belgium

Submitted to Wildlife biology
Abstract

Human-Wildlife Impacts (HWI) occur due to interactions between wildlife and human activities in our increasingly anthropogenic world and typically result in economic losses or increased health- and safety risks. HWI can be especially prevalent where urbanization encroaches upon natural areas, or in fragmented human-dominated landscapes. An example of such situation is the re-occurrence of wild boar in Flanders (northern Belgium). Flanders is one of the most densely populated areas of Europe and is characterized by a severely fragmented landscape. The recent return of wild boar to Flanders challenges managers to find solutions for a sustainable co-existence between humans and wild boar. As crop damage is increasing and targeting preventive measures efficient requires identifying high risk areas, we assessed the influence of the landscape around a field, as well as field-specific characteristics on the likelihood of wild boar crop damage. Because most of the reported damage in Flanders occurs in grasslands (cultivated to produce hay) and maize fields, we focused on these. We used boosted regression trees and the brglm-technique to construct distribution models explaining spatial patterns of crop damage. We found that for maize fields, landscape-level variables such as the proportion of maize, grassland, forest and urbanized areas in the surroundings of the field are key factors determining the probability of damage. In contrast, field-specific variables only played a minor role. For grasslands, both field-specific and landscape characteristics affected damage probability: a higher probability of damage was associated with decreasing distance to nearest forest, increasing distance to the nearest road, the use of inorganic fertilization and increasing age of the grassland. Our results suggest that the risk of crop damage by wild boar can potentially be mitigated by changes in agricultural practices that alter grassland characteristics, and by targeting preventive measures towards high risk maize in well-defined locations.

Key words: boosted regression trees, crop damage, damage probability, risk assessment, species distribution modelling, Sus scrofa
Introduction

The growth of the human population and increasing human-dominated land use challenges many wildlife species’ survival (McKee et al. 2004). At the same time, there are wildlife species which show increasing numbers due to protection or due to adjustment towards anthropogenic pressures or even benefit from urbanized areas through behavioural adaptations (Bateman and Fleming 2012, Lowry et al. 2013, Luniak 2004). Moreover, landscapes become increasingly anthropogenic and fragmented which causes wildlife to come more and more into contact with human activities (Barua et al. 2013, Messmer 2000). Human-Wildlife Impacts (HWI, here defined according to (Redpath et al. 2013) as the impacts due to the interactions between wildlife and humans or their activities) often involve an economic component. The economic component is often one of the main limiting factors in acceptance by stakeholders of wildlife (Carpenter et al. 2013). A better understanding of how, where and why HWI emerge is essential for wildlife managers to find viable solutions and thus decrease negative impacts and increase stakeholder acceptance of wildlife (Messmer 2009, Young et al. 2010). Incorporating animal ecology in a multidisciplinary approach of wildlife management can be valuable towards the search for such solutions in HWI (Fryxell et al. 2014).

Wild boar (*Sus scrofa* L.) is a suid native to Eurasia. Wild boar increasingly cause HWI (Massei et al. 2015) as since the 1960’s, populations throughout Europe and in other parts of the world where feral pigs have become an invasive alien species, started expanding (Massei et al. 2015, Mayer 2018, Saez-Royuela and Telleria 1986, Salvador and Fernandez 2018). Wild boar has become one of the most widespread mammals in the world (Keuling et al. 2018). More than half a century ago, wild boar disappeared in Flanders (Northern Belgium) due to overhunting. Wild boar re-appeared in 2006 in the Eastern province of Limburg in two geographically distinct locations. These founder populations were geographically not connected to existing populations, excluding natural recolonization by migration; however there is no confirmed
information on the origin of these founder populations. Wild boar in Flanders is regarded a native game species and the objectives for wild boar management are since 2016 set by stakeholder consultation for each out of 10 management zones (no presence allowed, zero acceptance of damage, limited damage as well as local populations allowed, but no further population increase). The current wild boar population in Flanders is still characterized by increasing population numbers and an expanding distribution range (Scheppers et al. 2014).

As Flanders is one of the most densely populated areas of Europe (FOD Economie 2011, Linell et al. 2001), characterized by a severely fragmented landscape, wild boar presence results in increasing incidence of HWI. HWI involving wild boar can include disease transmission, traffic collisions and damage to agricultural crops (Bieber and Ruf 2005, Morelle et al. 2016, Treves et al. 2006). Especially damage to agricultural land is a growing concern because of the high economic impact for individual farmers. Annually reported damage to crops ranges from hundred thousand euros in the Netherlands to more than twenty million euros in France (Carnis and Facchini 2012, Faunafonds 2014). In Flanders, the extent of crop damage from wild boar is largely unknown as there is no standardized monitoring method (Rutten et al. 2018). This lack of data does not allow to assess the actual extent nor the potential extent of crop damage with future wild boar expansion. As this is an essential aspect in a risk assessment involving stakeholder acceptance, we urgently need better insights into what attracts wild boar to specific fields or grasslands where they cause damage (Rutten et al. 2019a).

To obtain spatially explicit predictions of wild boar damage risk across Flanders, we applied Species Distribution Modelling (SDM). SDM tools are widely used to gain ecological insights into a species distribution and make predictions of a species’ (potential) distribution across landscapes (Elith and Leathwick 2009, Jiménez-Valverde and Lobo 2007). SDM are frequently applied as risk assessment tools (Acevedo et al. 2014, Jiménez-Valverde et al. 2011, Mateo-Tomás et al. 2012). SDM methods can not only be used to predict the distribution of the species itself (e.g. for wild boar (Morelle et al. 2016,
Saito et al. (2012)), but also to model the distribution of a species’ impact, such as crop damage by wildlife (Sitati et al. 2005, Tourenq et al. 2001). For wild boar, SDM-approaches have been used to predict damages to croplands in a rice-paddies dominated region in Japan by Saito et al. (2011) and in southern Italy by Ficetola et al. (2014). One of the central assumptions of SDM are that training data (i.e. the input data used to calibrate the model) are representative of the environmental conditions of the regions for which predictions are made (Elith et al. 2010). Therefore, we aimed to construct an SDM explaining crop damage patterns by wild boar based on data collected in the severely fragmented landscape of Flanders itself. Farmers reporting damage indicate they have the impression that field-specific factors like fertilization and the previous cultivated crop on a field can be key factors explaining differences in damage probability between neighbouring fields. Therefore, we did not only incorporated landscape variables but also field-specific factors that can be related to agricultural practices (i.e. fertilization, previous crop etc.). Finally, as wild boar can show a substantial plasticity in adjusting to human-dominated environments (Stillfried et al. 2017a), the Flemish landscape provides an interesting case study on damage patterns by wild boar in highly anthropogenic areas. Combining both landscape and field-specific aspects in a SDM we aim to answer following questions:  
- Which factors in the landscape attract wild boar to a specific field in a highly anthropogenic area?  
- Does landscape fragmentation affect damage patterns?  
- Do field-specific factors have an extra explanatory power additional to landscape factors?  
This study aims to increase our understanding about the characteristics of agricultural fields that are most likely to be damaged when wild boar are present based on landscape characteristics, with the ultimate goal to generate region-wide predictive crop damage risk maps. Moreover, a better understanding on field-specific characteristics allows to construct scenarios highlighting how relevant field-specific variables can modify landscape-related risks on crop damage.
Methods

Study area

Our study area encompasses the Flanders region of Northern Belgium. Flanders has a surface of 13 587 km² and has a cool temperate and moist climate (Metzger et al. 2013) with an annual average temperature of 9.7°C and 800 mm rainfall. Flanders has mainly a flat or gently undulating landscape from sea level in the West to 150 m above sea level in the South and East. The Flemish landscape is highly fragmented with only 11% forests, 53% agricultural land, 30% built-up areas and the remaining 6% consists of water, swamps, heathlands, natural grasslands, estuaries and dunes (Demolder et al. 2014). An intense intertwinement of natural, agricultural and urbanized areas is crossed by a dense road network (5.08 km/km², Vercayie and Herremans 2015). The current distribution area of wild boar is mainly limited to the eastern provinces of Limburg, Antwerp and Flemish Brabant but their distribution range is expanding towards the centre of the region (Figure 4.1). We selected a study area of approximately 1000 km² in Limburg, near the border to the Netherlands (Figure 4.1) where farmers often reported damage in an earlier survey (Rutten et al. 2019a).
Data collection

Farmers and hunters were contacted through the local farmers- and hunters-organizations and were asked to report damaged fields by phone between June 2015 and September 2018. Rooting damage, damage due to feeding, wallowing or damage after sowing of all possible crops (including grasslands which are further also classified as an agricultural crop) could be reported. During a field visit, we assessed if the field was indeed damaged by wild boar and the amount of damage was recorded during the field visit using a drone (Rutten et al. 2018). For the purpose of this study, the information on the amount of damage was however not used as we used a binary classification: damaged/not damaged. Field-specific characteristics were recorded (crop type, fertilization type, variety in case of maize field, age in case of grassland). During the field visit, the farmer or hunter was also asked to point out an undamaged field of the same crop, within 500 m from the damaged field. After controlling that this second field was indeed undamaged using a drone (Rutten et al. 2018), the same
information on field-specific characteristics was recorded as for damaged fields. The set of undamaged fields acts as a control group in this study, resulting in a presence-absence dataset to test landscape and field-specific factors. In the end, only maize fields and grasslands were included in this study since we only received 5 reports of other crops being damaged by wild boar.

In total, 275 fields were recorded between 2015 and 2018 (Figure 4.1, Table 4.1) of which 90 were grasslands and 185 maize fields.

<table>
<thead>
<tr>
<th></th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grasslands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undamaged</td>
<td>2</td>
<td>13</td>
<td>6</td>
<td>7</td>
<td>28</td>
</tr>
<tr>
<td>Damaged</td>
<td>3</td>
<td>41</td>
<td>16</td>
<td>2</td>
<td>62</td>
</tr>
<tr>
<td>Maize fields</td>
<td>39</td>
<td>70</td>
<td>46</td>
<td>30</td>
<td>185</td>
</tr>
<tr>
<td>Undamaged</td>
<td>9</td>
<td>29</td>
<td>13</td>
<td>22</td>
<td>73</td>
</tr>
<tr>
<td>Damaged</td>
<td>30</td>
<td>41</td>
<td>33</td>
<td>8</td>
<td>112</td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>124</td>
<td>68</td>
<td>39</td>
<td>275</td>
</tr>
</tbody>
</table>

Landscape variables

To identify landscape elements that influence wild boar crop damage patterns, we used a set of 15 landscape variables for each field of which data were collected during this study as well as for all fields in Flanders. This allowed us to make projections on crop damage probability for the rest of Flanders (see further) (Table 4.2). These calculations were done using the yearly Flemish parcel registration maps (EPR, from 2015 until 2017 (EPR of 2018 was not yet available), Flemish Land Agency), the land use map of Flanders NARA level 1 (Poelmans and Van Daele 2014) and the map of hunting grounds in which hunting rights are provided in Flanders (ANB, Agency for Nature of Forest). All calculations were conducted in ArcMap (version 10.4.1, ESRI inc.). A buffer zone was drawn around each field with a width of 1.25 km which was
used to calculate area-related variables. This radius of 1.25 km results in a
buffer area of at least 5 km² which corresponds to wild boar home ranges in
Wallonia (southern Belgium), that range between 4.9 km² (+/- 5.6 km²) for
males and 2.5 km² (+/- 3.74 km²) for females and juveniles (Prévot and Licoppe
2013).

As wild boar is often considered to be a forest-dwelling species (Briedermann
1990) and studies of Lombardini et al. (2016) and Ficetola et al. (2014) found
that shelter provided by forests contributed strongly to the risk that adjacent
fields were targeted by foraging wild boar, we calculated a number of landscape
variables related to forest: percentage of forest within the buffer, distance of
field edge to the nearest forest patch and percentage of the field edge which is
forest. As also areas with scrub and other low natural cover (natural grasslands,
heathlands, wetlands, fens, reeds etc.) are part of the natural habitat of wild boar
(Thurfjell et al. 2009), the percentage of this scrub and other low natural cover
in each buffer was calculated as well. Arable land was shown to be an important
factor driving crop damage by wild boar (Bleier et al. 2012), therefore we
calculated a set of landscape variables related to the presence of cropland and
grassland: percentage of agricultural land in each buffer, yearly percentage of
area maize fields within the buffers and yearly percentage of grasslands area in
each buffer for to the year in which the land was damaged (as the EPR of 2018
was not yet available, mean percentages from 2015 until 2017 were used for
damaged fields of 2018). The spatiotemporal behaviour of wild boar can change
in human-dominated landscapes (Podgórski et al. 2013): wild boar covered
twice as much distance on daily basis in urban areas although they have smaller
home ranges compared to primeval forests, there are less seasonal variations of
movements in urban areas and urban areas wild boar were mainly nocturnal. To
assess potential effect of behavioural adjustments on damage patterns, the
following variables related to urbanization were calculated: percentage of
urbanized area (defined as industrial areas, build-up areas, infrastructures and
commercial areas) in the buffers and the distance from the field until the nearest
urbanized area. Because we also wanted to assess the possible role of
fragmentation of natural habitats on crop damage patterns, variables such as
road density (total length of primary and secondary roads divided by the buffer area) and distance to nearest road were calculated. Considering forest together with scrub and other low natural cover as habitat, we moreover calculated the habitat patch density (number of habitat patches divided by the buffer area and mean habitat patch size in each buffer. Lastly, to include hunting activity, we calculated the distance to the nearest hunting area and the percentage of hunted ground within each buffer based on the hunting grounds map. If the field in question itself was situated within a hunting area, the distance to the nearest hunting area was set to zero.

Table 4.2: Landscape variables calculated for the observed damaged and undamaged fields in the collected dataset and for all fields in Flanders with both mean and standard error.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dataset</th>
<th>Flanders</th>
<th>Variable</th>
<th>Dataset</th>
<th>Flanders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of edge which is forest</td>
<td>25.07 ±2.63</td>
<td>5.41 ±0.024</td>
<td>Percentage of urbanized area in buffer</td>
<td>17.71 ±0.70</td>
<td>25.52 ±0.018</td>
</tr>
<tr>
<td>Distance until nearest forest (m)</td>
<td>52.62 ±6.74</td>
<td>163.02 ±0.26</td>
<td>Percentage of forest in buffer</td>
<td>23.29 ±0.96</td>
<td>7.00 ±0.013</td>
</tr>
<tr>
<td>Distance until nearest road (m)</td>
<td>610.08 ±34.48</td>
<td>394.17 ±0.56</td>
<td>Percentage of agriculture in buffer</td>
<td>51.07 ±1.20</td>
<td>62.43 ±0.026</td>
</tr>
<tr>
<td>Distance until nearest urbanized area (m)</td>
<td>48.46 ±5.22</td>
<td>18.41 ±0.071</td>
<td>Percentage of maize in buffer</td>
<td>16.13 ±0.59</td>
<td>15.80 ±0.011</td>
</tr>
<tr>
<td>Distance until nearest hunting area (m)</td>
<td>11.48 ±4.06</td>
<td>9.98 ±0.16</td>
<td>Percentage of grasslands in buffer</td>
<td>19.46 ±0.53</td>
<td>19.75 ±0.015</td>
</tr>
<tr>
<td>Road density (length of roads divided by buffer area)</td>
<td>3.43 ±0.16</td>
<td>5.67 ±0.0049</td>
<td>Percentage of hunted area in buffer</td>
<td>85.56 ±1.16</td>
<td>94.12 ±0.018</td>
</tr>
<tr>
<td>Habitat Patch Density in Buffer</td>
<td>9.63 ±0.25</td>
<td>7.47 ±0.0080</td>
<td>Percentage of low cover in buffer</td>
<td>2.04 ±0.13</td>
<td>1.25 ±0.0030</td>
</tr>
<tr>
<td>Mean Habitat Patch Size in buffer (km²)</td>
<td>0.034 ±0.0022</td>
<td>0.012 ±0.00036</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Field-specific variables

For each reported agricultural field, field-specific variables were collected during the field visits (Table 4.3). Fertilizer application was subdivided into two separate variables: organic and inorganic fertilizer use. For maize fields, the specific variety was characterized using the precocity-value (FAO-value from the Food and Agricultural Organization) defining the timing of ripening of the variety. For grasslands, the age of grasslands was recorded (grasslands older than 5 years are considered permanent grasslands). As farmers had the impression that remains of maize from the previous year are rooted up by wild boar in grasslands, a binary variable representing the crop of the previous year (maize or no maize (i.e. grass, cereals, potatoes or nothing)) was included.

Table 4.3: Field-specific variables collected for the observed damaged and undamaged maize fields and grasslands by wild boar in the study area in Flanders with their categories or with its mean and standard error.

<table>
<thead>
<tr>
<th>Maize fields</th>
<th>Grasslands</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organic fertilization</td>
<td>Organic fertilization</td>
</tr>
<tr>
<td>(Categories)</td>
<td>(Categories)</td>
</tr>
<tr>
<td>Inorganic fertilization</td>
<td>Inorganic fertilization</td>
</tr>
<tr>
<td>(Categories)</td>
<td>(Categories)</td>
</tr>
<tr>
<td>FAO-value for precocity</td>
<td>Age</td>
</tr>
<tr>
<td>243±54</td>
<td>4.59 ± 2.20</td>
</tr>
<tr>
<td>Crop of previous year</td>
<td>Maize/No</td>
</tr>
<tr>
<td>(Categories)</td>
<td>maize</td>
</tr>
</tbody>
</table>

Distribution models

Variables were screened for multicollinearity using the Pearson’s correlation coefficient (R-value) with R = 0.7 as a threshold to remove correlated variables (Dormann et al. 2013). Variables which showed multicollinearity with the largest number of other variables were removed step by step until none of the remaining variables showed a R > 0.7.

Models were constructed separately for maize fields and grasslands. For each crop, 3 models were built: a landscape model including only landscape variables, a field-specific model including only field-specific factors and a
combined model that included all variables marked as relevant predictors of wild boar damage by the best landscape and best field-specific model (see below).

We use Boosted Regression Trees (BRT) in R (R Core Team 2016) as SDM algorithm as they typically have a high predictive performance, do not have a need for prior data transformations or for elimination of outliers, automatically incorporate interaction effects and are able to fit nonlinear relationships (Elith et al. 2008). We used the dismo R-package (version 1.1-4, Hijmans et al. 2017) to develop the BRT models. For BRT-modelling, optimal parameter settings (learning rate (lr), tree complexity (tc) and bag fraction (bf)) were first determined. To define optimal settings, the model was run with a range of possible settings of lr, tc and bf, averaging model outputs over 10 iterations to minimize variability between runs. Optimal settings were selected based on a minimum model output of 1000 trees and a minimal predicted deviance. Once the optimal settings were derived, the least important variables were removed by refitting the model each time with the removal of one variable while assessing the change in predictive deviance according to the procedure of Elith et al. (2008). The final model is obtained when the change in predictive deviance exceeds the original standard error of the full model (Elith et al. 2008). Additionally, the AUC-value (area under the receiver-operating characteristics curve or roc-curve) of the final model was calculated; AUC-values vary between 0 and 1 with values higher than 0.5 reflecting a better ability for a model to discriminate between damaged and undamaged fields than by random chance.

The available dataset on wild boar damage in grasslands was substantially smaller than for maize field (90 versus 185, Table 4.1), and resulted in model fitting issues when running the BRTs for grasslands, as the model was not able to reliably discriminate between patches with and without boar damage. We therefore opted for grasslands for an alternative regression technique, namely ‘BrGLMs’ (Bias reduction in Binomial-Response Generalized Linear Models) using the brglm R-package (version 0.6.1, Kosmidis 2017). Here, relevant
‘robust’ predictor variables were selected using a model selection procedure based on Akaike’s information Criterion (AIC of the MuMIn R-package version 1.42.1, Barton 2018) in which AIC-values of all possible models (reflecting all possibilities of variable combinations) are first calculated. The relative importance of variables was then determined by summing the AIC-weights of all models in which a specific variable was included (Rouffaer et al. 2017). Robust variables are indicated by high AIC-weights (>0.5) and model-averaged estimates which are higher than their standard error (Burnham and Anderson 2002). Non-robust variables were removed from the model. 10-fold cross validation was used for the final set of robust variables in which the BrGLM is run 10 times, each time withholding randomly 10% of the data. Model estimates and AUC-values were then averaged over these 10 runs.

Projection of crop damage probability

We developed maps indicating damage probability for each field in Flanders, representing damage probability under the condition that wild boar would be present (as wild boar presence itself is not modelled here) and the crop in question would be cultivated at the specific field. Subsequently, we ran three model scenarios for both maize fields and grassland. For the first run, projections are only based on the landscape model, thus not incorporating field-specific variables, showing the spatial distribution of damage risk in Flanders due to landscape-level characteristics only (see above). To simulate a ‘worst case’ scenario, the combined model including both landscape and field-specific variables was run whereby the field-specific variables were set such that they result in the highest damage risk (high risk scenario). A ‘best case’ scenario was then made with all field-specific variables in the combined model set such that they result in the lowest damage risk possible (low risk scenario). These high and low risk scenarios reflect the influence of field-specific variables on the damage probability extent and thereby the possible impact of changes in field-specific agricultural practices that farmers can implement or can allow farmers to consider crop planning changes.
Extrapolation outside the training range of the dataset of a SDM can return less reliable results (Fitzpatrick and Hargrove 2009). To quantify the degree of extrapolation in our projections, the extent of environmental differences between model training and projection data is calculated using multivariate environmental similarity surface (MESS) maps using the `ecospat` package (version 3.0, Broennimann et al. 2018). MESS-analysis measure the similarity between the dataset used to train the model and the newly projected areas. Positive MESS-values reflect that the full range of the new variable values are included in the original dataset, while negative MESS-values reflect variable conditions which are not included in the training data, thereby identifying areas where the model is extrapolating.

**Results**

*Maize fields*

Due to multicollinearity, the variables percentage of agricultural cover, percentage of urban cover and mean habitat patch size were removed from the dataset. Model selection of the landscape model resulted in the further removal of distance to nearest urbanized area, distance to nearest forest patch, distance to nearest road, road density and habitat patch density. The final landscape model for maize therefore consequently includes the percentage of maize, -grassland, -forest, -urban area, -hunting area, and of scrub and other low natural cover within each buffer and the percentage of forested edge of the field (Table 4.4). For the field-specific maize model, model selection resulted in the removal of organic fertilization resulting in a final model including the precocity-value and inorganic fertilization.
Table 4.4: Final landscape model, field-specific model and combined model with remaining variables for maize fields to explain crop damage probability by wild boar in Flanders. Variable importance (%) and model parameters (mean total deviance, training data correlation and AUC-value (area under the curve value)) are shown.

<table>
<thead>
<tr>
<th>Variable importance</th>
<th>Landscape model</th>
<th>Field-specific model</th>
<th>Combined model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage Maize</td>
<td>16.32</td>
<td>/</td>
<td>15.29</td>
</tr>
<tr>
<td>Percentage Grassland</td>
<td>16.89</td>
<td>/</td>
<td>17.23</td>
</tr>
<tr>
<td>Percentage Forest</td>
<td>18.20</td>
<td>/</td>
<td>14.26</td>
</tr>
<tr>
<td>Percentage Urban area</td>
<td>16.50</td>
<td>/</td>
<td>15.29</td>
</tr>
<tr>
<td>Percentage Hunting area</td>
<td>11.57</td>
<td>/</td>
<td>10.94</td>
</tr>
<tr>
<td>Percentage Forested edge</td>
<td>10.62</td>
<td>/</td>
<td>8.90</td>
</tr>
<tr>
<td>Percentage Low cover</td>
<td>10.90</td>
<td>/</td>
<td>11.66</td>
</tr>
<tr>
<td>FAO-value for precocity</td>
<td>/</td>
<td>81.66</td>
<td>4.19</td>
</tr>
<tr>
<td>Inorganic fertilization</td>
<td>/</td>
<td>18.34</td>
<td>2.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>Landscape model</th>
<th>Field-specific model</th>
<th>Combined model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean total deviance</td>
<td>1.34</td>
<td>1.37</td>
<td>1.37</td>
</tr>
<tr>
<td>Training data correlation</td>
<td>0.81</td>
<td>0.40</td>
<td>0.83</td>
</tr>
<tr>
<td>AUC-value</td>
<td>0.97</td>
<td>0.73</td>
<td>0.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model settings</th>
<th>Landscape model</th>
<th>Field-specific model</th>
<th>Combined model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning rate</td>
<td>0.0005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>Bag fraction</td>
<td>0.67</td>
<td>0.50</td>
<td>0.75</td>
</tr>
<tr>
<td>Tree complexity</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Number of trees</td>
<td>9750</td>
<td>2650</td>
<td>1450</td>
</tr>
</tbody>
</table>

The landscape model was the most accurate model (AUC-value of 0.97) explaining crop damage probability with percentage of maize, -grassland, -forest and -urban area in buffers as the most predictive variables (Table 4.4). Based on the relationship between the variables and damage probability (Supplementary materials, Appendix 4.1), the following patterns can be found: damage probability increases with increasing area forest surrounding a field although when the percentage forested area surrounding the field increases over 40%, damage probability does not increase further. Lowest damage probabilities occur when urbanization around the field is approximately 10%, percent, lower as well as higher percentages of urbanization results in higher damage risk. The more maize is available around a field, the lower the chance
that the field will be damaged. This effect also stabilizes around 30% of the area being covered by maize. Furthermore, when the area around a field exists for more than 30% out of grasslands, damage probability decreases for a particular grassland. When >90% of the area around the field exists out of hunting grounds as well as when less than 15% of the field edge exists out of forest, damage probability decreases. Finally, lowest damage probabilities are found when the area around a field has no scrub and other low natural cover but damage probability slightly increases when there is scrub and other low natural cover surrounding the field.

The field-specific model showed a lower accuracy (AUC of 0.73) with the precocity-value being the most predictive value. The combined model has a similar accuracy as the landscape model (AUC of 0.96) (Table 4.4) but field-specific variables only contribute for a small part (6.44%) to the predictive power.

**Grasslands**

Because of multicollinearity, the variables percentage of agricultural land, urbanized area and mean habitat patch size in buffers were deleted from the dataset. The final landscape grassland model includes distance to nearest forest patch and distance to nearest road as the only robust variables (Table 4.5, Supplementary materials, Appendix 4.2).

The combined model shows the highest-AUC value (AUC of 0.91), thus the combination of field-specific- and landscape variables explains damage probability of grasslands the best. The grassland model shows a positive effect of the use of organic fertilization and a negative effect of the use of inorganic fertilization on grassland damage probability. Increased age was found to result in a higher damage probability. Moreover, when maize was cultivated before, the damage risk in the first year increases. Furthermore, damage probability decreases when distance to nearest forest patch increases and distance to road decreases.
<table>
<thead>
<tr>
<th></th>
<th>Landscape model</th>
<th>Field-specific model</th>
<th>Combined model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>p-value</td>
<td>Estimate</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.0± 0.17</td>
<td>0.00076±0.00087</td>
<td>2.31± 0.32</td>
</tr>
<tr>
<td>Distance to nearest forest</td>
<td>-1.022±0.13</td>
<td>0.0065±0.0028</td>
<td>-0.3±0.20</td>
</tr>
<tr>
<td>Distance to nearest road</td>
<td>1.050±0.23</td>
<td>0.017±0.015</td>
<td>1.57±0.036</td>
</tr>
<tr>
<td>Age grassland</td>
<td>-0.52±0.18</td>
<td>0.18±0.078</td>
<td>-0.48±0.16</td>
</tr>
<tr>
<td>Use of organic fertilization</td>
<td>-1.89±0.14</td>
<td>0.019±0.0060</td>
<td>-2.25±0.25</td>
</tr>
<tr>
<td>No maize as previous crop</td>
<td>-3.02±0.52</td>
<td>0.050±0.0046</td>
<td>-5.73±0.63</td>
</tr>
<tr>
<td>Use of Inorganic fertilization</td>
<td>2.06±0.20</td>
<td>0.010±0.0048</td>
<td>2.80±0.21</td>
</tr>
<tr>
<td>No maize as previous crop:Age</td>
<td>0.96±0.20</td>
<td>0.033±0.023</td>
<td>1.40±0.19</td>
</tr>
<tr>
<td>Model parameters</td>
<td>AUC</td>
<td>0.74 ± 0.016</td>
<td>AUC</td>
</tr>
</tbody>
</table>
Model projections to Flanders

Figure 4.2: a) Projection on crop damage probability by wild boar for Flanders based on the landscape model for maize fields if on all fields maize was cultivated and wild boar are present and b) the density distribution of damage probability in Flanders of a high risk and low risk scenario of the combined model compared with the distribution based on the landscape model.
Figure 4.3: a) Projection on crop damage probability by wild boar for Flanders based on the landscape model for grasslands if on all fields grasslands were cultivated and b) the density distribution of damage probability in Flanders of a high risk and low risk scenario of the combined model compared with the distribution based on the landscape model.
Damage probability based on the landscape model of maize fields showed a heterogeneous distribution of damage probability in Flanders with a generally lower damage probability in the West compared to the East (Figure 4.2a). Around urbanized areas, there is generally a higher damage probability although model extrapolation occurs in these region (see MESS maps, Supplementary materials, Appendix 4.3). As expected due to the limited contribution of field specific variables to the total model, the high- and low risk scenarios show only a limited change in damage probability distribution and extent compared to the landscape model (Figure 4.2b).

Damage probability based on the landscape model for grasslands shows an overall higher damage probability in Flanders compared to maize fields and less geographic variations (Figure 4.3a). The high- and low risk scenarios show a large change in damage probability distribution and extent (Figure 4.3b) compared to the landscape model reflecting the importance of field-specific characteristics. In general, for estimating grassland damage probability, model extrapolation did occur more compared to maize (MESS maps, Supplementary materials, Appendix 4.3) indicating that projections on wild boar damage to grasslands based on the landscape model need to be interpreted more cautiously.

**Discussion**

With an increasing number of Human Wildlife Impacts (HWI) due to damage to agricultural land in Flanders, there was an urgent need to better understand factors attracting wild boar to specific fields. Flanders is strongly divergent from the areas where modelling studies on wild boar have been done in the past, as the region is characterized by severe fragmentation and human-dominated land-use. Behavioural flexibility typically allows wild boars to adapt to regional habitat conditions, and therefore, bespoke distribution models for characterizing crop damage probability in such highly fragmented landscape needed to be developed. We gained knowledge on the influence of both
landscape-level and field-specific factors in maize fields and grasslands. Landscape characteristics were found to have a more profound effect compared to field-specific characteristics when predicting damage probability for maize fields but for grasslands, a combination of field-specific characteristics and landscape characteristics determine the damage probability. Simultaneous increase of wild boar population and urban areas in the last decades results in more frequent presence of wild boar around urban areas. With Flanders being one of the most extremely anthropogenic areas in Europe, our gained insights will be of use for other urbanized areas in Europe where wild boar are becoming more successful.

As forest gives shelter to wild boar (Lombardini et al. 2016), the positive effect of forest on increasing damage probability in maize fields and from shorter distance to forest for grasslands was not surprising, and confirms results of previous studies: Ficetola et al. (2014) found increased crop damage risk from wild boar with intermediate forest cover values in a mosaic of agricultural and natural areas in Italy, Saito et al. (2011) showed that crop damage risk in rice paddies in Japan increased with shorter distances from forest edges, a study in Germany identified the majority of rooting in grasslands close to woodlands (Daim 2015) and a study in a nature reserve in Flanders found increased rooting risk in natural grasslands surrounded by forest (Volckaert 2013). Moreover, as scrub and other low natural cover can also provide shelter opportunities, this can explain the increase in damage risk when scrub and other low natural cover increases around a field. An increasing percentage of grasslands in the direct surroundings of maize fields results in less shelter, which can explain the negative correlation between damage probability for maize fields and the presences of grasslands. Although maize also provides (seasonal) cover to wild boar (Schley et al. 2008), the positive correlation between damage probability and more maize cover can be explained by the balance between supply and demand of maize as a food source in the direct surroundings: the more maize is cultivated in the environment, the lower the damage probability of an individual field due to sufficient availability of maize as a food source.
Although landscape fragmentation has been found to negatively influence wild boar occurrences (Virgós 2002), wild boar is also identified as a species which shows a substantial plasticity in adjusting to human-dominated environments (Stillfried et al. 2017a) by adjusting their spatiotemporal behaviour (i.e. urban wild boar show smaller home ranges but increasing travel distance to meet energetic requirements compared to rural wild boar (Podgórski et al. 2013)). As Flanders has an extremely fragmented landscape and is one of the most densely populated areas of Europe, Flanders presented us an interesting case study to see if there are limits in wild boars’ plasticity towards landscape fragmentation inhibiting crop damage probability. However, we did not find any effects of fragmentation per se on crop damage probability: compared to the other factors we evaluated, variables characterizing fragmentation did not seem to affect the crop damage probability of a maize field or grassland, thus suggesting that wild boar movements while foraging do not seem to be inhibited by landscape fragmentation. This adds to our findings of a landscape genetic analysis of the recolonization of wild boar in Flanders, which also showed that wild boar expansion in Flanders did not seem to be negatively affected by severe landscape fragmentation (Rutten et al. 2019b).

We only found an effect of hunting once the of the area around a field is covered by hunting grounds exceeds 90%, damage probability starts to decrease in maize fields. However, as information of hunting bag records is only collected at the level of municipalities and game management units but no exact locations of hunting records were available, we could not define hunting pressure around a specific field. Moreover, as it has been shown that the effect of hunting on spatial activity varies between hunting methods ranging from smaller home ranges to increased flee distances (Keuling et al. 2008c), we suggest to further investigate the mechanisms between hunting pressure, hunting methods and damage patterns before implementing or adjusting hunting strategies to decrease crop damage risk.

Field-specific factors play only a minor role in comparison to landscape factors for maize fields. This is in contrast to grasslands were fertilization, age of the
grasslands and the crop of the previous year play an important role according to our grassland model. While the use of organic fertilization was found to have an inhibiting effect on damage probability for grasslands, using inorganic fertilization increases damage probability. Macro invertebrates like earthworms and grubs are part of the diet of wild boar (Barrios-Garcia and Ballari 2012) and macro invertebrate communities are affected by fertilization (Laznik and Trdan 2014, Potter et al. 1985). We thus assume that the effect of fertilization on crop damage probability could be explained by the effect of fertilization on this dietary source from wild boar: Laznik and Trdan (2014) found an increasing weight of earthworms in the soil with increasing K$_2$O-values and an increasing number of grubs with increasing P$_2$O$_5$ content, both being frequent components of inorganic fertilization. We thus expect this explains increasing rates of grassland rooting as wild boar root up macro invertebrates from the soil. Also positive effects of the use of organic fertilizers on increasing earthworms have previously been found (Daim 2015, Whalen et al. 1998), which could also result in more rooting to find earthworms, something we did not find in our results. As earthworm populations are also influenced by factors like soil temperature and moisture which have not been included in this study, the dynamics of earthworm populations in agro-ecosystems are expected to be more complex and also are generally known to be poorly understood (Whalen et al. 1998). Furthermore, grassland damage probability increased with increasing grassland age. This with the exception when maize was previously cultivated, in these cases damage probability to grasslands was higher for young grasslands (0-1 year). As intense management of grasslands leads to decreased soil macro invertebrate biodiversity (Bardgett and Cook 1998), older grasslands in which soil biota can develop with time results in a more diverse dietary source which has been found to be preferred by wild boar (Bueno et al. 2009).

Making projections based on the maize- and grassland landscape models results in projection maps of the potential risk and geographical distribution of crop damage risk in Flanders, under condition wild boar would recolonize the whole region. In maize fields, a generally higher damage probability was found in the East compared to the West. However, we want to point out that, due to a larger
percentage of the area used for agriculture in the West of the region, the overall economic impact of wild boar crop damage could still be higher in the West compared to the East. Furthermore, although agricultural fields around large urban areas seem to have a high damage probability, it should be noted that it is also in these regions that our distribution models are most strongly extrapolating beyond the range of model training conditions. As the training set of the model originates from the current distribution area of wild boar, which is not yet as urbanized as the most extremely urbanized regions in Flanders, these projections around highly urbanized areas are less reliable. Furthermore, the high- and low risk scenarios reflect the influence of field-specific variables on damage probability and to which extent the adjustment of these variables could lead to mitigating damage risk. The difference between high- and low risk scenarios were low for maize field projections due to the limited influence of field-specific factors of crop damage probability. However, we found large differences in the overall extent of damage risk under the different scenarios for grasslands. The high risk scenario showed an overall extreme high damage risk in Flanders while the low risk scenario showed an overall low damage risk in grasslands. This reflects the large influence of field-specific variables compared to the influence of landscape variables, thus highlighting the potential high impact of agricultural management decisions on grassland damage by wild boar. For maize fields, adjusting agricultural practices to affect field-specific damage risk will be less effective although here, the implementation of preventive measures (i.e. electric fences) can now be targeted more efficient as well as targeted crop planning can allow to grow maize in low damage risk fields and grow other crops in high damage risk fields (if maize would be planted).

Our current landscape-based risk maps assume the presence of wild boar in the whole area, currently not being the reality as the species is still expanding its range. An important next step in order to further improve the risk maps in order to guide prioritization of preventive measures will be to integrate the probability of wild boar presence in the near future due to expansion. A previous study using a landscape genetic approach to assess recolonization of
wild boar in Flanders indicated that wild boar expansion was hardly inhibited by fragmentation but is mainly driven by the presence of forested areas (Rutten et al. 2019b). Therefore it is expected that the future wild boar expansion and distribution will not be random. A detailed assessment of further wild boar distribution will be essential to define the future damage risk in Flanders.

Because of multicollinearity, some variables were not included in the models although they might have equally important contributions towards explaining crop damage patterns as those that were included in the model building. However, the goal of this model is to act as a predictive tool for problem-solving and our damage probability projections represent the current, best explanatory power of where wild boar damage will occur in highly fragmented landscapes. The variables chosen here for modelling are directly related to the species’ known ecology, increasing the transparency and understandability of our results – and these are important criteria when developing any statistical model (Starfield 1997). We are therefore convinced that our modelling and projection results on wild boar crop damage patterns in maize fields and grasslands increased our understanding of where crop damage is most likely to occur. Our research now allows to guide preventive actions (e.g. placing electric fencing on fields) more efficient and cost-effective, thereby minimizing HWI’s in a highly fragmented and anthropogenic landscape. Moreover, farmers may use our findings for crop planning strategies aimed at minimizing damage risks.
Management implications

Damage management is an important part of wild boar management. Understanding the impact of certain agricultural practices like selection the crop, fertilization, crop variety selection etc. on damage risk can help farmers to decide on crop planning strategies or implement preventive measures while mitigating crop damage risk. This will reduce HWI’s by wild boar and consequently increase stakeholder acceptance of wildlife (Messmer 2009, Young et al. 2010). Combining the results of this study with possible further wild boar distribution maps allows in to generate maps showing further risks of wild boar damage to agricultural land. This risk mapping can guide decision taking on taking preventive measures or to developing effective management strategies to limit HWI’s (Acevedo et al. 2007, Červinka et al. 2015, Fischer et al. 2015, Linkov et al. 2006).

Acknowledgements

A. Rutten held a PhD grant (IWT.141377) from the Agency Flanders Innovation & Entrepreneurship (VLAIO), Flanders, Belgium. We thank all farmers for reporting damaged fields, Boerenbond and Hubertus Vereniging Vlaanderen (HVV) for their cooperation and contacts with the farmers and hunters. We thank Joery Van Looy and Jesse Daneels for assisting field work and preliminary data-analysis.
Chapter 5

Analysing the recolonization of a highly fragmented landscape by wild boar using a landscape genetic approach

Rutten, Anneleen\textsuperscript{1,2}; Cox, Karen\textsuperscript{2}; Scheppers, Thomas\textsuperscript{2}; Bram Vanden Broecke\textsuperscript{1}; Leirs, Herwig\textsuperscript{1}; Casaer, Jim\textsuperscript{2}

\textsuperscript{1}Evolutionary Ecology Group, Department of Biology, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium

\textsuperscript{2}Research Institute for Nature and Forest (INBO), Havenlaan 88 bus 73, B-1000 Brussels, Belgium

Published in Wildlife Biology, 2019, 2019 (1): 1-11. DOI: 10.2981/wlb.00542
Abstract

Wild boar are currently one of the most widespread mammals of the world and in many regions populations keep expanding. In Flanders (Belgium), the wild boar has returned since 2006 after almost half a century of absence and numbers are increasing fast. The Flemish landscape is severely fragmented and is one of the most densely populated areas in the world. Understanding the relationship between landscape structures and species biology is the basis of landscape ecology and increases the understanding of factors driving habitat use, recolonization and expansion. We conducted a landscape genetics study to identify factors driving wild boar expansion in Flanders. A total of 838 DNA-samples collected from the wild boar hunting bag between 2007 and 2016 were genotyped for 140 single nucleotide polymorphisms (SNPs). We show that the wild boar population expansion started from 2 local gene pools while staying relatively genetically distinct, though with some admixture. A third gene pool emerged around 2015 in the Northwest coming from the Netherlands and Germany. The landscape genetic analysis revealed that the main factors explaining the spatial genetic pattern are isolation by distance and forest cover which influenced gene flow positively. Forest fragmentation had no significant effect on genetic distances. As Human-Wildlife Impacts are increasing in line with wild boar expanding distribution range, understanding factors driving expansion during recolonization is essential for assessing the future dispersal of wild boar in Flanders. With a better insight in future dispersal, it will be possible to conduct risk assessments which target more efficient management actions to limit Human-Wildlife Impacts.

Key words: Single Nucleotide Polymorphisms, recolonization, fragmentation, expansion, landscape genetics, Sus scrofa
Introduction

The relationship between landscape structures and the biology of a species forms the basis of landscape ecology (Coulon et al. 2006). In this matter, landscape genetics has become an important tool to increase our understanding of landscape elements that affect diffusion. How a landscape is perceived by a certain species does not necessarily match human presumptions of landscape connectivity or habitat quality (Cushman et al. 2006). Understanding the interaction between geographical/environmental features and genetic variation can reveal discontinuities in the landscape such as barriers, the extent of landscape connectivity (Balkenhol et al. 2013, Cox et al. 2017, Kierepa and Latch 2016, Norman et al. 2016, Parks et al. 2015, Villedéy et al. 2016), aid understanding of colonisation and expansion processes of invasive species (Fischer et al. 2017) and delineate management units for species management (e.g. ungulates; Coulon et al. 2006, Frantz et al. 2009). These applications have all used landscape genetic approaches (Holderegger and Wagner 2008, Manel et al. 2003).

Knowledge about landscape elements that affect dispersal is essential for developing effective management strategies of fast expanding species (Fischer et al. 2017, Frantz et al. 2010, Nikolov et al. 2009, Scandura et al. 2008). One such fast expanding species is the wild boar (*Sus scrofa* L.). Since the 1960s wild boar populations have been expanding throughout Europe (Acevedo et al. 2007, Massei et al. 2015, Saez-Royuela and Telleria 1986) and much of their invasive distribution range (New Zealand: Bengsen et al. 2018, Australia: Choquenot et al. 1996, North America: Mayer 2018, & McClure et al. 2015, South America: Salvador and Fernandez 2018). This makes wild boar currently one of the most widespread mammals in the world found in all continents apart from Antarctica (Keuling et al. 2018).

Flanders (northern Belgium) is one of the most densely populated areas in Europe, with a human population density of 462 persons per square kilometre
Chapter 5

(report FOD Economie 2011, Linell et al. 2001). An intense interweaving of small natural areas, forest remnants and agricultural areas, interspersed by a dense road network (5.2 km/km²) results in a highly fragmented landscape. Wild boar populations have been present and/or expanding their distribution area during the last decades in all neighbouring regions and countries around Flanders (Carnis and Facchini 2012; Jansman et al. 2013; Massei et al. 2015; Morelle et al. 2016; Sodeikat and Pohlmeyer 2003; Widar 2011). After an absence of more than fifty years, wild boar reappeared in Flanders in 2006 at two geographically distinct locations from where range expansion within Flanders began. There is no official information on where these population originated from and these populations were geographically not connected to neighbouring populations. Despite this highly fragmented landscape, wild boar numbers in Flanders are increasing rapidly (Scheppers et al. 2014). At the same time, the anthropogenic landscape causes frequent contacts between wild boar and human activities resulting in and emerging Human-Wildlife Impacts (HWI, here defined as impacts due to interactions between wildlife and humans activities according to (Redpath et al. 2013)) (Young et al. 2010).

As wild boar numbers and Human-Wildlife Impacts increase in Flanders (Rutten et al. 2018), a risk assessment is urgently needed. Essential in a risk assessment are future dispersal predictions, its consequences and effective management strategies. Effective strategies to prevent wild boar impacts should be based on the understanding of factors driving population increase and colonisation of new areas by wild boar (Massei et al. 2015, Veličković et al. 2016). Reforestation, agricultural intensification and climate change have been found to be the main drivers of wild boar population growth (Massei et al. 2015, Saez-Royuela and Telleria 1986). However, patterns related to range expansions are poorly understood (Morelle et al. 2016). Forest coverage has recently been found to influence wild boar dispersal (Morelle et al. 2016) but the influence of landscape fragmentation, characterizing the Flemish landscape, is currently unknown.
In this study we assess the recolonization and expansion processes of wild boar in Flanders using a landscape genetic analysis. We aim to increase our understanding of genetic connectivity in a severely fragmented landscape. We analysed a set of DNA samples, collected from the Flemish hunting bag since the beginning of the reappearance of wild boar, to (i) assess the evolution in genetic population structure during recolonization and (ii) understand the role of landscape connectivity throughout recolonization. We expected to find that fragmentation has affected wild boar dispersal during the recolonization in the past decade as its spatiotemporal behaviour and occurrence patterns have been found to be affected by fragmentation (Podgórski et al. 2013, Virgós 2002). This information about factors driving recolonization is crucial to be able to predict future dispersal of wild boar in a fragmented landscape like Flanders and will be essential for the development of a risk assessment to guide effective management strategies.

Methods

Study area

The study area largely encompasses the current distribution area of wild boar in Flanders. Presently, wild boar are almost exclusively found in the eastern part of Flanders (Figure 5.1). Hunting records show that recolonization started at two distinct geographical areas in the eastern province of Limburg (Scheppers et al. 2014).
Figure 5.1: Dashed areas: distribution area of wild boar in Flanders (northern Belgium) in 2017. Yellow: distribution area in 2007 at start of recolonization. As the exact location of DNA-samples is unknown, the red dots represent the centroid coordinates of the municipality within the specific game management unit of each sample.

Sample collection

Tissue samples of the lower jaw of wild boar have been collected since the early beginning of the recolonization from hunting bag and road kills. No animals were killed for this study. From the available set of approximately 2000 tissue samples a set of 838 samples was selected for the period from 2007 until 2016 (Table 5.1, Figure 5.1). The selection aimed at an even spread of samples across the distribution area per year. Samples are not evenly distributed over the years given the limited number of wild boar shot in the beginning of the recolonization. The exact coordinates of these samples were not available as hunters are not obliged to report the exact coordinates of hunted wild boar. However all samples could be assigned to a specific municipality and a specific game management unit (GMU). The geographic unit used in the landscape genetic analysis is the part of a municipality located within a certain GMU.
Table 5.1: Overview of all samples selected for this study from 2007 until 2016 and the number of samples selected for the landscape genetic analysis.

<table>
<thead>
<tr>
<th>Year</th>
<th>Total number of samples</th>
<th>Number of samples used in landscape genetic analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>2008</td>
<td>23</td>
<td>6</td>
</tr>
<tr>
<td>2009</td>
<td>50</td>
<td>15</td>
</tr>
<tr>
<td>2010</td>
<td>30</td>
<td>6</td>
</tr>
<tr>
<td>2011</td>
<td>53</td>
<td>23</td>
</tr>
<tr>
<td>2012</td>
<td>106</td>
<td>34</td>
</tr>
<tr>
<td>2013</td>
<td>104</td>
<td>41</td>
</tr>
<tr>
<td>2014</td>
<td>146</td>
<td>26</td>
</tr>
<tr>
<td>2015</td>
<td>274</td>
<td>111</td>
</tr>
<tr>
<td>2016</td>
<td>35</td>
<td>23</td>
</tr>
<tr>
<td>TOTAL</td>
<td>838</td>
<td>286</td>
</tr>
</tbody>
</table>

**DNA extraction and genotyping**

DNA was extracted from the 838 tissue-samples using the Qiagen DNAEasy Blood and Tissue kit (QIAGEN Inc). Samples were genotyped for 150 single nucleotide polymorphisms (SNPs) using the Illumina porcine SNP60 genotyping beadchip (Ramos et al. 2009). These SNPs were selected out of the set of 351 SNPs used by Goedbloed et al. (2013) in an analysis of population structure of Northwest European wild boar and that are known to be polymorphic for wild boar in this area (Goedbloed et al. 2013). The set of 150 SNPs was selected based on observed heterozygosity (<0.8) and highest minor allele frequency (MAF). SNPs were genotyped by LGC Genomics (LGC Group, Hoddesdon, UK) using their KASP (Kompetitive Allele Specific PCR) genotyping assay (He et al. 2014). A maximum limit of 5% missing values was used to exclude SNPs because of low genotyping quality.
Hardy-Weinberg equilibrium and linkage disequilibrium

As deviation from Hardy-Weinberg equilibrium (HWE) or linkage disequilibrium (LD) can have a substantial impact on population structure analysis (Waples 2015), HWE and LD were assessed while taking into account potential population structure. Clusters were defined using Discriminant Analysis of Principal Components (DAPC) with the adegenet R package (version 2.0.0, Jombart and Collins 2015) in R (Version 3.5.2, R Core Team 2016). The absence of assumptions concerning HWE or LD make DAPC a useful method to detect a potential structure (Jombart et al. 2010). Tests were performed for each identified genetic cluster. HWE was tested using the adegenet R package, LD was tested using the genetics R package (version 1.3.8.1, Warnes et al. 2015) and results were corrected for multiple testing using the Bonferroni correction.

Population structure during expansion

Using the Bayesian clustering approach in Structure (version 2.3.4, Pritchard et al. 2000), we conducted clustering of all 838 samples in subsequent three-year sections (to have a sufficient number of samples in each section), each with an overlap of one year (Table 5.2). For each temporal section, we used 10 independent runs with a burn-in of 100 000 Markov chain Monte Carlo (MCMC) iterations and 1 000 000 sampling iterations for each of 1 to 8 potential clusters (K) assuming correlated allele frequencies and admixture with a variable alpha value of 1/K. Variable alpha values were used to correct for small sample sizes (Wang 2017). The optimal value of K for each section was determined using the method of Evanno et al. (2005) embedded in Structure Harvester (webversion v0.6.94, Earl and VonHoldt 2012). Average q-values (proportion of each individual’s ancestry from population K) of all 10 runs of the chosen K were calculated in Clumpp (version 1.1.2, Jakobsson and Rosenberg 2007). We assigned individuals to a cluster based on their highest average q-value to a certain population K to determine the evolution of clusters throughout the years.
Table 5.2: Overlapping section per 3 years and set of Goedbloed et al. (2013) used for the Structure analysis.

<table>
<thead>
<tr>
<th>Section</th>
<th>Years</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2007-2009</td>
<td>90</td>
</tr>
<tr>
<td>2</td>
<td>2009-2011 + Netherlands/Germany set</td>
<td>335 (133 + 202)</td>
</tr>
<tr>
<td>3</td>
<td>2011-2013</td>
<td>263</td>
</tr>
<tr>
<td>4</td>
<td>2013-2015</td>
<td>524</td>
</tr>
<tr>
<td>5</td>
<td>2015-2016</td>
<td>309</td>
</tr>
</tbody>
</table>

We included samples genotyped for the same set of SNPs as used by Goedbloed et al. (2013), which were collected between 2008 and 2010 in the Netherlands and eastern Germany. For this comparison all Dutch and German samples were selected within a buffer of 75 km around Flanders. A limit of 75 km was set to assess potential geographic connections with the Flemish population clusters. The selection resulted in 202 samples of Goedbloed et al. (2013) and was analysed together with the corresponding 2009-2011 section of our data. A separate Structure analysis was run including all 838 Flemish + 202 Dutch and German samples of Goedbloed et al. (2013) to assess connections of clusters over the years.

Of the final genetic clusters identified, observed heterozygosity ($H_o$) and expected heterozygosity ($H_e$) per section were calculated using the adegenet R package. Between inferred clusters, pairwise fixation index ($F_{st}$) as a measure for differentiation among populations, was calculated using the hierfstat R package (version 0.04-22, Goudet and Jombart 2015). Temporal changes were tested using simple linear models (lm function in R) in which year is used as response variable and $H_o$, $H_e$ or $F_{st}$ are used as explanatory variables.
Landscape genetics analysis

a) Resolution

To identify landscape elements that could potentially influence wild boar dispersal, a landscape genetic analysis was conducted. As each sample was assigned to a specific area (the part of a municipality within a specific GMU) with different sizes (median = 43 km², min. size = 1 km² and max. size = 114 km²), a subset of samples with a maximum size of 40 km² was selected. Segelbacher et al. (2010) showed that the effects of landscape features in gene flow should be studied at the scale as large as the movement distance of the studied species. Dispersal capacity of wild boars in Wallonia (southern Belgium) is found to vary between 2.49 ± 3.74 km (mean ± standard deviation) for juveniles and females and 4.90 ± 5.65 km for males (Prévot and Licoppe 2013), resulting in a potential movement area of 20 to 75 km². (when supposing a circle buffer with a radius of 2.49 to 4.9 km). However, taking into account the large variation in home ranges which have been reported ranging between 0.68 and 48.3 km² (Garza et al. 2017), we had to find a balance between a useful scale to conduct a landscape genetic analysis taking into account dispersal- and home ranges and geographic representation of our samples. The selection of maximum 40 km² area size was therefore found the optimal balance taking all these aspects into account. This resulted in a subset of 286 samples out of the 838 Flemish samples used in this landscape genetic analysis (Table 5.1).

Spatial genetic patterns respond to changes in the landscape structure (Landgut et al. 2010) but as the Flemish landscape did not alter considerably in the last decade, we do not consider the time span of 10 years of these genetic data long enough to be influenced by recent landscape changes. Therefore, we did not use temporal restrictions for sample selection (Landgut et al. 2010).
b) Individual pairwise genetic distances

We used Rousset’s a genetic distance (GD) (Rousset 2000) as it was shown to be among the most accurate metrics for landscape genetic approaches (Shirk et al. 2017a). Individual pairwise Rousset’s a genetic distances between 286 samples were calculated using Spagedi (version 1.5, Hardy and Vekemans 2002).

As hunting records showed geographically distinct founder populations (see results), and we want to focus this analysis on mechanisms driving recolonization, the pairwise genetic distance of these 286 individuals which show a high ancestry to different population clusters, and therefore individuals who are potentially from different historical founder populations and are thus not related, are not calculated. This way we get a better insight in the mechanisms driving recolonization within clusters and drivers of potential admixture between the source-populations.

To define which individual pairwise genetic distances were not calculated, Structure results were used: these represented pairwise combinations of which both individuals show a q-value higher than 0.8 to different populations clusters (for example: pairwise genetic distance between individual 1 with a q-value of 0.90 to population 1 and individual 2 with a q-value of 0.90 to population 2 is not calculated while the pairwise genetic distance between this individual 1 and an admixed individual 3 with a q-value of 0.45 to population 2 is calculated).

c) Interpatch pairwise genetic distances

We calculated the mean interpatch genetic distances between all 33 municipality-GMU areas arising from the individual pairwise combinations. This resulted in a patch-based landscape genetic approach with a total of 501 mean interpatch pairwise genetic distances (it should be noted that in theory there are 528 possible interpatch combinations, but this number decreased due to reduction in the number of pairwise combinations in step b). The patch-based approach accounts for potential dependency among individuals from the same patch.
d) Random point buffers

As the exact geographical coordinates of samples within each area are unknown, we randomly selected a point location in each area. After connecting the random points by straight lines a buffer with a radius of 3.5 km was drawn around each line (Figure 5.2). A buffer radius of 3.5 km results in a buffer surface of 40 km² around each location, corresponding to the threshold of a maximum surface area of 40 km² used to select samples for the landscape genetic analysis (step a). In total, 501 buffers corresponding to the 501 interpatch pairwise genetic distances were drawn. To account for variability of the landscape within each area, we repeated the random selection of a point location 100 times after which the buffers were drawn for each of the 100 times and landscape variables (next step e) were calculated for each of the 100 sets of corresponding buffers (Figure 5.2).

Figure 5.2: Workflow of landscape genetic analysis with visualization of step d-e (determining landscape variables as in figure e urbanized coverage in buffers).
Chapter 5

\textbf{\textit{e)} Landscape variables}

Within the buffers, a set of 10 landscape variables (Table 5.3) were calculated using the CORINE land cover classes (inventory of 2012 (EEA 2012); Supplementary materials, Appendix 5.1). Forest cover was calculated by merging all forest classes in CORINE and was included in this analysis as wild boar are a forest-dwelling species (Briedermann 1990) and forest coverage has been shown to influence wild boar dispersal (Morelle et al. 2016). The CORINE classes of grasslands, heathlands, wetlands etc. were grouped into low natural cover and were incorporated as these landscape types are also part of the natural habitat of wild boar (Thurfjell et al. 2009). Because agricultural crops are an important food source for wild boar (Schley and Roper 2003), the percentage of arable land use (agricultural coverage) was included as a third variable. Finally, the percentage of urbanized land was used as a forth land cover type.

Isolation by distance (IBD, the increase in genetic differentiation among individuals with geographic distance) can affect genetic distances due to limited dispersal when distances between individuals increases (Frantz et al. 2010), the Euclidian distance (km) between the random chosen locations of each polygon was calculated.

A set of fragmentation measures for each buffer was calculated: road density (the length of primary, secondary roads and motorways (OpenStreetMap 2018) divided by buffer area), forest patch density (number of forest patches divided by buffer area), mean forest patch size (forest cover area divided by number of forest patches), mean forest patch edge-area ratio (forest patch perimeter divided by patch area) and mean nearest neighbour forest patch distance (distance from each forest patch to the nearest other forest patch).
Table 5.3: Overview of landscape variables used in the landscape genetic analysis. Mean values and standard deviation (SD) of these variables have been calculated over all 501 buffers and over the 100 different sets of buffers. These variables were selection based on knowledge of habitat use of wild boar and to assess potential effects of forest fragmentation on wild boar dispersal.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest coverage</td>
<td>30.46 %</td>
<td>8.20 %</td>
</tr>
<tr>
<td>Low nature coverage</td>
<td>9.33 %</td>
<td>5.64 %</td>
</tr>
<tr>
<td>Agricultural coverage</td>
<td>33.03 %</td>
<td>13.24 %</td>
</tr>
<tr>
<td>Urbanization coverage</td>
<td>26.15 %</td>
<td>7.52 %</td>
</tr>
<tr>
<td>Euclidian distance</td>
<td>23.45 km</td>
<td>13.67 km</td>
</tr>
<tr>
<td>Road density</td>
<td>10.67 km/ha</td>
<td>2.27 km/ha</td>
</tr>
<tr>
<td>Natural forest patch density</td>
<td>0.12 patches/buffer</td>
<td>0.036 patches/buffer area</td>
</tr>
<tr>
<td>Mean natural forest patch size</td>
<td>0.55 ha</td>
<td>0.22 ha</td>
</tr>
<tr>
<td>Mean forest patch edge-area ratio</td>
<td>0.062 m/m²</td>
<td>0.38 m/m²</td>
</tr>
<tr>
<td>Mean nearest neighbour forest patch distance</td>
<td>256.17 m</td>
<td>77.73 m</td>
</tr>
</tbody>
</table>

f) Model

For each of the 100 generated datasets including the 10 landscape variables and genetic distances, we set up linear regression models with the 501 interpatch pairwise genetic distances as response variable, using maximum-likelihood population effects parameterization (MLPE) (Clarke et al. 2002) as it was shown that these models perform well in landscape genetic regression methods (Shirk et al. 2017b). The R package nlme was used to fit all models (version 3.1-137, Pinheiro et al. 2016). In order to define the correlation structure that accounts for non-independence of pairwise distances, the R package corMLPE (version 0.0.2, Pope 2018) was used.

The 10 landscape variables were standardized around the mean and were screened for multicollinearity using Variance Inflation Factors (VIF) applying a threshold of VIF < 3 to remove variables (Fox and Monette 1992). Model optimization using backwards selection was performed based on Akaike’s
information criterion (AIC) with a threshold of AIC-difference of 2 (Burnham and Anderson 2004).

Using the AIC-values of each run, weighted AIC-scores were calculated according to Wagenmakers and Farrell (2004) using the qpcR R package (version 1.4-1, Spiess 2018) coefficient results were averaged over the 100 model runs using these weighted AIC-scores and the number of significant runs per coefficient was calculated.

Results

Preliminary analysis

Ten out of 150 SNPs were excluded because of too high percentage of missing values, resulting in 140 SNPs for data analysis.

Using the Bayesian Information Criterion (BIC) in function of the number of clusters (k-value) (Supplementary materials, Appendix 5.2), DAPC analysis resulted in 4 clusters. Twelve out of 140 SNPs did show deviations from HWE in at least 3 out of 4 clusters. Structure and landscape genetic analysis were run with the total set of 140 SNPs and were compared with the reduced set of 128 SNPs. Deviations from LD and HWE were found to have a substantial impact on population structure analysis as outcomes differed between both SNP sets (results not shown). We therefore used the reduced 128 SNP set for further analysis (Supplementary materials, Appendix 5.3).
Population structure during expansion

The *Structure* analysis of the three-year sections detected two distinct clusters in the beginning of the recolonization in 2007-2009 (Figure 5.3): a more eastern population (EP) and a more western population (WP), although there was evidence of some admixture between the two clusters (Supplementary materials, Appendix 5.4). In the following years (2009-2011, 2011-2013 and 2013-2015) the two clusters stay largely geographically delineated (Figure 5.3) although there remains a certain degree of admixture (Supplementary materials, Appendix 5.4). The analysis of the section 2015-2016, revealed a new third cluster in the Northwest (NWP). This NWP cluster was found to be related to the samples of the Netherlands and Germany as shown in the separate *Structure* analysis including all samples of Flanders, the Netherlands and Germany (Figure 5.3). Moreover, the comparison with the samples of the Netherlands and Germany shows a connection of the western population (WP) to southern samples from Germany in the 2009-2011 section, however this connection is not clear when all samples were analysed together.

Changes in $H_e$, $H_o$ and $F_{st}$ over the years were not found to be significant (Table 5.4). Differentiation between NWP and WP was lower than between NWP and EP. DAPC clustering results (4 clusters) were compared with *Structure* clustering results (3 clusters: Supplementary materials, Appendix 5.2): clustering patterns are mainly similar, the fourth DAPC-clusters is largely a subdivision of the NWP cluster.
Figure 5.3: STRUCTURE analysis of wild boar in Flanders per 3 years during expansion and of the total dataset including Dutch and German samples (Goedbloed et al. 2013). Pie chart size varies with the number of individuals per location and show the portion of individuals assigned to each population (based on highest q-values): red represent the eastern population (EP), blue represent the western population (WP), yellow represent north-western population (NWP) which is present in 2009-2011 in the Netherlands and Germany and which emerges in Flanders starting from 2015.
Table 5.4: Trend of observed Heterozygosity ($H_o$), expected Heterozygosity ($H_e$) per Structure cluster and pairwise population differentiation ($F_{st}$) between Structure clusters (significant $F_{st}$ values are marked with * if 95% confidence intervals did not include 0) over the years.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of clusters</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_o$ EP</td>
<td>0.36±0.0029</td>
<td>0.36±0.0028</td>
<td>0.35±0.00018</td>
<td>0.36±0.00028</td>
<td>0.36±0.00066</td>
<td>0.95</td>
</tr>
<tr>
<td>$H_o$ WP</td>
<td>0.36±0.00081</td>
<td>0.36±0.0003</td>
<td>0.35±0.00027</td>
<td>0.36±0.00022</td>
<td>0.39±0.0035</td>
<td>0.26</td>
</tr>
<tr>
<td>$H_o$ NWP</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>0.37±0.001</td>
<td></td>
</tr>
<tr>
<td>$H_e$ EP</td>
<td>0.43±0.000078</td>
<td>0.43±0.0012</td>
<td>0.41±0.0007</td>
<td>0.40±0.00098</td>
<td>0.41±0.000015</td>
<td>0.86</td>
</tr>
<tr>
<td>$H_e$ WP</td>
<td>0.43±0.00056</td>
<td>0.41±0.00013</td>
<td>0.39±0.000074</td>
<td>0.44±0.000063</td>
<td>0.36±0.00081</td>
<td>0.98</td>
</tr>
<tr>
<td>$H_e$ NWP</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>0.37±0.001</td>
<td></td>
</tr>
<tr>
<td>$F_{st}$ WP-EP</td>
<td>0.0064*</td>
<td>0.027*</td>
<td>0.0052*</td>
<td>0.0050*</td>
<td>0.035*</td>
<td>0.52</td>
</tr>
<tr>
<td>$F_{st}$ WP-NWP</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>0.010*</td>
<td></td>
</tr>
<tr>
<td>$F_{st}$ EP-NWP</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>0.050*</td>
<td></td>
</tr>
</tbody>
</table>
Landscape genetics analysis

In total, 27,316 individual pairwise genetic distances were used to calculate 501 mean interpatch genetic distances. Of all 10 landscape variables that were considered, agricultural coverage, mean forest patch size and road density were excluded because of high VIF, removing multicollinearity between variables. Backwards model selection did not result in further exclusion of variables resulting in a final model including forest coverage, urbanized coverage, low nature coverage, Euclidean distance, forest patch density, mean forest patch edge-area ratio and mean nearest neighbour forest patch distance.

Table 5.5: Model parameters of final model. Coefficient averages (Coeff.), Standard Error (SE), t-values and p-values are averaged over the 100 model runs. Variable effect is considered significant with a p-value ≤ 0.05.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff.</th>
<th>SE</th>
<th>t-value</th>
<th>p-value</th>
<th>Proportion significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.0069</td>
<td>0.17</td>
<td>0.040</td>
<td>0.97</td>
<td>0</td>
</tr>
<tr>
<td>Forest Coverage</td>
<td>-0.22</td>
<td>0.040</td>
<td>-5.48</td>
<td>&lt;0.0001</td>
<td>100</td>
</tr>
<tr>
<td>Urban Coverage</td>
<td>0.016</td>
<td>0.042</td>
<td>0.38</td>
<td>0.59</td>
<td>0</td>
</tr>
<tr>
<td>Low Nature Coverage</td>
<td>-0.046</td>
<td>0.034</td>
<td>-1.35</td>
<td>0.26</td>
<td>16</td>
</tr>
<tr>
<td>Euclidean Distance</td>
<td>0.37</td>
<td>0.028</td>
<td>13.25</td>
<td>&lt;0.0001</td>
<td>100</td>
</tr>
<tr>
<td>Forest Patch Density</td>
<td>-0.0079</td>
<td>0.033</td>
<td>-0.24</td>
<td>0.52</td>
<td>3</td>
</tr>
<tr>
<td>Mean forest patch edge-area ratio</td>
<td>-0.00073</td>
<td>0.019</td>
<td>-0.045</td>
<td>0.51</td>
<td>5</td>
</tr>
<tr>
<td>Mean nearest neighbour forest patch distance</td>
<td>0.012</td>
<td>0.030</td>
<td>0.42</td>
<td>0.46</td>
<td>4</td>
</tr>
</tbody>
</table>

Both forest coverage and Euclidean distance were found to influence genetic distance as all 100 runs showed a significant effect of these variables. Forest coverage has a negative effect on genetic distance and Euclidean distance was
found to have a positive relation towards genetic distance (Table 5.5 & Supplementary materials, Appendix 5.5). Although model optimization resulted in the inclusion of urban coverage, low nature coverage, forest patch density, mean forest patch edge-area ratio and mean nearest neighbour forest patch distance, these were mainly non-significant in the full model (few or no runs were significant for these variables).

**Discussion**

Knowledge about the influence of landscape composition and fragmentation on species distribution is essential for the design of effective risk assessments to reduce Human-Wildlife Impacts, particularly in anthropogenically disturbed landscapes such as Flanders. The population structure analysis gave us an insight in the first 10 years of recolonization: dispersal did not happen at random and a specific population structure was found. Flemish wild boar populations expanded from two separated local gene pools, confirming the presence of the two geographically distinct areas identified at the start of the expansion (Scheppers et al. 2014). Although there was quite some admixture between populations and little genetic differentiation was found, population differentiation was significant indicating that gene flow was influenced by environmental factors, slowing down the formation of one panmictic population.

As there is no official information on where the Flemish populations originated from, the finding that some clusters were related to populations from neighbouring regions gives insight into the potential origin of the Flemish population clusters. By assessing the Flemish population structure together with samples from the Netherlands and Germany, we found that a third gene pool which emerged in 2015 in Northwest Flanders is related to samples from the Netherlands and Germany suggesting that natural migration from these neighbouring countries has occurred since 2015. Additionally, the western Flemish population shows some similarities to the samples from southern
Germany in the yearly sections, however, as this connection was not found in the population structure of all samples together we cannot say with certainty that southern Germany would indeed be a source of this population in Flanders. Moreover, no connection between the East Flemish population and any of the other sampled populations was found. Previous analysis of wild boar ancestry in the Netherlands, Belgium and Germany using microsatellites (Jansman et al. 2013) did not show any relatedness between the Flemish population and other neighbouring populations (including eastern Walloon populations). The research of Breyne et al. (2014) on genetic profile of Flemish wild boar also failed to find a clustering connection between Flemish and the sampled Walloon populations. Possible explanations for the absence of clear origin of all Flemish populations in neighbouring populations can be the lack of samples from other potential neighbouring origin-clusters as we did not have access to samples from southern Belgium, potential reintroductions from non-neighbouring populations or a lack of clear relatedness with other clusters due to founder effects (Broders et al. 1999). Further research including DNA-samples from other populations would be needed to identify origin populations.

The landscape genetic analysis gave us insight in the connectivity of the highly fragmented landscape in Flanders. The results showed that increasing forest cover is linked to decreasing genetic distance, which is not surprising as even in urban areas it has been shown that wild boar has a strong preference for natural landscapes (Stillfried et al. 2017). The importance of forest during expansion was also found in the studies of Morelle et al. (2016) and Saito et al. (2012). Clear patterns of isolation by distance (IBD) were also found as Euclidean distance was positively related to genetic distance, confirming findings of Frantz et al. (2012) and Renner et al. (2016). We did not find any significant effects of fragmentation in contrast to what we expected, as wild boar spatiotemporal behaviour and number of occurrences have been found to be affected by fragmentation (Podgórski et al. 2013; Virgó 2002). However, it has also been shown that wild boar show substantial plasticity regarding adjustment to human-dominated environments i.e. landscape of fear (Stillfried et al. 2017); wild boar tolerate human presence by modulating their risk perception and even
use human-associated habitat classes. Wild boar dispersal is thus not limited by human-dominated environments and they may not experience a negative effect of habitat fragmentation during expansion. Another explanation could be that the lack of effects on fragmentation is the consequence of hunting-induced dispersal. It has been shown that hunting can influence spatial utilisation, increase flight distances or home ranges although effects differ with different hunting methods and sometimes no change or decreasing home-ranges are found (Boitani et al. 1994, Calenge et al. 2002, Keuling et al. 2008c). We could not test this potential effect of hunting-induced dispersal as we did not have detailed information on variation in hunting pressure through the study time-span and study area.

Landscape genetic studies benefit from an individual-based approach, often using datasets with specific coordinates of sample locations (Cushman et al. 2006, Holderegger and Wagner 2008, Kristensen et al. 2018, Parks et al. 2015, Segelbacher et al. 2010). As we did not have the exact coordinates from each sample, we opted to conduct the landscape genetic analysis using a subset of samples which could be allocated to an area of a municipality within a GMU of less than 40 km² and used repeated random point locations in these areas to determine the influence of landscape variables. Working without exact coordinates, we expected that this would significantly reduce statistical power to connect landscape variability to genetic distances compared to working with specific coordinates. However, the landscape analysis resulted in interesting findings showing that with minimal time and cost effort (working with data which is available), it is possible to gain essential knowledge on factors driving expansion during recolonization. We are therefore convinced that the presented method is robust and results are solid.
Management implications

Human-Wildlife Impacts are the main limiting factor in stakeholders’ tolerance towards wildlife (Carpenter et al. 2013). The lack of an effect of habitat fragmentation on wild boar dispersal is likely to lead to future dispersal and range expansion in the highly fragmented landscape of Flanders. Conducting a risk assessment, in which the extent of future dispersal is crucial, can lead to developing effective management strategies to limit Human-Wildlife Impacts. The knowledge gained through this landscape genetic analysis will now allow us to incorporate the influence of landscape elements in species distribution modelling (SDM) approaches and thereby improving predictions for the future wild boar distribution in Flanders.

Acknowledgements

We would like to thank Sabrina Neyrinck and Nancy Van Liefferinge for lab analysis and Axel Neukermans for tissue sampling. We want to thank dr. Daniel Goedbloed for his dataset on wild boar in the Netherlands and Germany. We want to express our appreciation to dr. Michel Haagdorens and dr. Luci Kirkpatrick for their valuable advice on the manuscript. Finally, we want thank the hunters for tissue sampling and the ANB (Agency for nature and forests) for information on hunting bag data.
Additional remarks in the dissertation after publication of the paper

Although the term dispersal is used throughout this paper, the term expansion of the distribution area may be a more correct term. Dispersal is generally used to describe individual movements to a new place for settling and breeding but that is not what we assessed in this analysis. Therefore, expansion which focuses on colonisation processes (of which dispersal is of course a part) is regarded to be a more correct term. Moreover, in the methods section where we mention dispersal capacity of wild boars in Wallonia, natal dispersal distance might have been a more correct term as this is expressed as a distance and not as a flux (distance/time).

Furthermore, we used the term “Northwest Flanders” to indicate the distribution of certain clusters which might be confusing as these terms were based on the current distribution range but not based on the whole of Flanders. Similarly, the term southern Germany was used based on the distribution of the available samples of Germany but not on the whole of Germany. Focussing on the whole of Flanders and Germany might imply more correct terms as northern part of the province of Antwerp and West Germany.

Finally we mention in the discussion that we are convinced that the presented method is robust and results are solid. We want to clarify that we did not imply that the method as such is robust but that the method resulted in robust results making the method robust for this type of analysis.
Future distribution of wild boar in a highly anthropogenic landscape: models combining hunting bag and citizen science data

Rutten, Anneleen¹,²; Casaer, Jim²; Swinnen, Kristijn RR³; Herremans, Marc³; Leirs, Herwig¹

¹ Evolutionary Ecology group (EVECO), University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium

² Research Institute for Nature and Forest (INBO), Havenlaan 88 bus 73, B-1000 Brussels, Belgium

³ Natuurpunt Studie, Coxiestraat 11, B-2800 Mechelen, Belgium

Abstract

Wild boar is one of the most widespread mammals of the world and in many regions wild boar populations continue to expand. Especially in highly anthropogenic landscapes, increasing numbers of wild boar lead to a rising number of contacts with human activities causing Human-Wildlife Impacts. In the heavily fragmented landscape of Flanders (northern Belgium) where the wild boar re-appeared in 2006 after more than half a century of absence, it is crucial to get a better understanding of the probable further distribution of wild boar in order to assess potential impacts in the near future. Wild boar occurrences have been collected by two citizen science programs: through an online observation platform and based on the reported locations of wild boar shot by hunters. This allowed us to construct a MaxEnt habitat suitability model. We constructed a new approach to define background manipulation to correct for sampling bias due to uneven sampling effort or due to areas in which hunting is not allowed based on the construction of bias files using this information. Model outcomes based on this new approach for background manipulation were compared with the known method of spatial thinning. All model outcomes were found comparable reflecting the utility of our new approach when limited data are available and spatial thinning would result in insufficient data for modelling. Our MaxEnt models show that coniferous forest, deciduous forest, maize, scrub and other low cover play a key role in increasing the habitat suitability for wild boar. Built up areas and the extent of habitat diversity only had a minor influence on habitat suitability reflecting wild boars’ behavioural flexibility to adapt to human-dominated landscapes. Unoccupied suitable habitat is mainly found in the centre of Flanders, although highly scattered. Habitat suitability in the West of Flanders was limited.

**Key words:** *Sus scrofa*, MaxEnt, Species distribution modelling, citizen science, sampling bias
Introduction

Landscares are becoming increasingly anthropogenic and fragmented, causing wildlife to come more into contact with human activities (Barua et al. 2013, Messmer 2000). Human-Wildlife Impacts (HWI, here defined according to Redpath et al. (2013) as impacts due to interactions between wildlife and human activities) are the main limiting factors in acceptance of wildlife by stakeholders (Carpenter et al. 2013). Wild boar (Sus scrofa L.) related HWI are increasing since the 1960’s when wild boar populations started to expand and increase throughout the original native range in Europe and in other parts of the world where feral wild boar are non-native (Massei et al. 2015, Mayer 2018, Saez-Royuela and Telleria 1986). These population expansions resulted in wild boar becoming one of the most widespread mammals in the world (Keuling et al. 2018). HWI involving wild boar include damage to agricultural crops, traffic collisions and disease transmission (Bieber and Ruf 2005, Morelle et al. 2016, Treves et al. 2006).

In Flanders (northern Belgium), wild boar disappeared after the second world war due to overhunting. However, since 2006, after more than half a century of absence, wild boar re-emerged in several locations. In the eastern province of Limburg they reappeared in two geographically distinct locations. These founder populations were geographically not connected to populations abroad excluding natural recolonization by migration; however there is no confirmed information on the origin of these populations (Rutten et al. 2019b). Since their return, both population numbers and distribution ranges are increasing but the distribution is currently still mainly limited to the Northeast of Flanders (Scheppers et al. 2014). During the last decennia, the Flemish landscape altered substantially due to economic growth, urbanization and agricultural intensification. Currently Flanders is one of the most densely human populated areas of Europe (Linell et al. 2001), characterized by a severely fragmented landscape and an intense intertwinement of agricultural, natural and urban areas. As a consequence wild boar presence results in an increasing numbers of HWI. By getting a better understanding of factors determining habitat
suitability for wild boar, the habitat suitability of currently uncolonized areas can be estimated, returning crucial information for conducting a risk assessment related to future potential wild boar expansion. Conducting such risk assessments allows to develop effective management strategies in order to avoid HWI’s (Červinka et al. 2015, Fischer et al. 2015).

Distribution models play an ever increasing role in conservation planning, wildlife management and related decision-making (Araújo and Guisan 2006, Kozak et al. 2008, Warren et al. 2014). Species distribution models (SDMs also called Habitat suitability models (HSM, Bellamy et al. 2013) or Ecological niche models (ENM, Sillero 2011)) relate species occurrences to environmental variables, thereby creating insights into habitat suitability for the species in question (Elith and Leathwick 2009). Characterizing the distribution of species does not only provide ecological insight but also allows to predict distribution across space and/or time (Elith and Leathwick 2009). These predictions may concern effects of climate change (Khanum et al. 2013), invasive species distribution potential (Bradley et al. 2010) or recolonization of native species (Swinnen et al. 2017). The use of models already showed its utility in wildlife management in the past (Baldwin 2009, Saito et al. 2012).

SDMs should be based on the understanding of the species biology, ecology and impact of human disturbance (Araújo and Guisan 2006). Various studies assess habitat suitability for wild boar in Europe (Bosch et al. 2014b, ENETwild consortium et al. 2019, Morelle et al. 2016). These studies mainly report forest to be among the most important factors determining habitat suitability. Furthermore, agriculture was found an important factor enhancing habitat suitability due to (seasonal) food availability and shelter opportunities (Herrero et al. 2006). Wild boar also prefer the presence of water in their home ranges (Ilse and Hellgren 1995). Wild boar show a substantial behavioural plasticity in adjusting to human-dominated environments (Stillfried et al. 2017a) and even became habituated to metropolitan areas like Barcelona and Berlin (Cahill et al. 2012, Kotulski and König 2008). However, their spatio-temporal behaviour has been found to be affected by human presence
Chapter 6

(Podgórski et al. 2013). Human interference and open vegetation without shelter opportunities have been reported to negatively impact the suitability of an area (Alexander et al. 2016, Bosch et al. 2014a). Moreover, even in urban area, natural food sources are, when available, preferred over anthropogenic food sources such as garbage (Stillfried et al. 2017b). Previous landscape genetic studies in Flanders and Wallonia showed no clear effects of forest fragmentation or fragmentation due to roads on landscape connectivity for wild boar (Dellicour et al. 2019, Rutten et al. 2019b) although more continuous, less fragmented landscapes are suggested to be more suitable (ENETwild consortium et al. 2019). The Flemish landscape is one of the most extremely fragmented and anthropogenic areas in Europe. Therefore, Flanders provides an interesting case study for understanding wild boar expansion mechanisms in extremely fragmented landscapes and to assess if there are limits towards wild boars’ behavioural plasticity in adjusting to anthropogenic pressures.

Species occurrence data are essential in SDM. Citizen science data are numerous and have been shown to advance knowledge of species occurrences and their distributions (Bonney et al. 2009), e.g. as a basis for SDM’s (Crall et al. 2015, Mair et al. 2017, Swinnen et al. 2017). In Flanders, wild boar observations of the data portal www.waarnemingen.be of Natuurpunt (Swinnen et al. 2018) provides a dataset that can be used in a SDM. Another citizen science data source of wild boar occurrences comes from Flemish hunters. Hunting bag data have already been suggested to be useful in SDM, however they are often recorded at too low resolution (ENETwild consortium et al. 2019). Given that Flemish hunters have the possibility since 2016 to register the exact coordinates of the location where individual wild boar were shot, a SDM approach at a detailed resolution is possible with these data.

The first aim of this research is to use SDM to gain a better understanding of the factors influencing habitat suitability and therefore determine the potential future distribution range of wild boar in the highly anthropogenic landscape of Flanders. The results from this research forms an essential element in risk
assessments to evaluate potential impacts when wild boar expand further into this severely fragmented and anthropogenic landscape.

Sampling bias can however affect the results of SDM’s (Guillera-Arroita et al. 2015): this bias can be related to uneven search effort, reporting behaviour or variation in detectability of the target species across the landscape. This can lead to localities that are biased in environmental space due to spatial autocorrelation of recorded species occurrences (Boria et al. 2014). By correcting for sampling bias, model over-fitting is avoided. There are some proposed methods which can be used to correct for sampling bias (e.g. spatial thinning or background manipulation with bias files (Boria et al. 2014, Kramer-Schadt et al. 2013)). This background manipulation using bias files is often based on occurrence data of similar taxa. Our datasets allowed a new approach to define bias. The online portal of waarnemingen.be contains observations of many species; using the observation characteristics of wild boar observers, we designed a method to define search effort. Moreover, as we have detailed maps of hunting grounds, we are able to assess sampling bias in hunting bag data due to the fact that hunting does not take place everywhere. Both search effort as hunting effort (based on hunting grounds) lead to the development of bias files to adjust background selection so sampling bias is taking into account. Based on this new approach defining sampling bias, as a second goal of this study we will assess how well our two approaches perform compared to spatial thinning to correct for sampling bias. With increasing number of citizen science data collections, our new approach can be of importance to take sampling bias into account when using this kind of data for modelling purposes.

**Methods**

**Study area**

Flanders has a surface of 13,587 km² and has a cool temperate and moist climate (Metzger et al. 2013) with an annual average temperature of 9.7°C and 800 mm rainfall. Flanders has mainly a flat or gently undulating landscape from
sea level in the West to 150 m above sea level in the South and East. The Flemish landscape is highly fragmented with only 11% forests, 53% agricultural land, 30% build up areas and the remaining 6% consists of water, swamps, heathlands, natural grasslands, estuaries and dunes (Demolder et al. 2014). An intense intertwinement of natural, agricultural and urbanized areas is crossed by a dense road network (5.08 km/km², Vercayie and Herremans 2015). The current distribution area of wild boar is mainly limited to the eastern provinces of Limburg, Antwerp and Flemish Brabant but their distribution range is expanding towards the centre (Figure 6.1). Current neighbouring wild boar populations are found in the Netherlands near the Belgian border and in Wallonia (Rutten et al. 2019b).

Figure 6.1: Study area of Flanders (grey) with 1383 occurrence records from waarnemingen.be (orange) and 1510 locations from hunting records in the eastern provinces (blue) and current distribution range of wild boar (dashed area, based on hunting records at municipal scale for 2018). The red polygon is the minimum convex polygon encompassing the data records used for MaxEnt background selection.
Data collection

Citizen science data on wild boar presences in Flanders was obtained from two different sources. Firstly the website waarnemingen.be, a portal of citizen scientist’s records of plant- and animal species in Belgium which started in 2008, containing over 33 million observations. The goal of waarnemingen.be is to be the digital notebook of all nature observations for users. The collected biodiversity information is shared with the public and species specific maps and statistics are reported. The data is used to gain information on species occurrences (Steeman et al. 2017), to monitor biodiversity or is used in species specific research projects (e.g. Swinnen et al. 2017). Records include sightings, footprints, rooting- and other foraging tracks (camera traps and records of road kills are not used as these include a different search effort). Of the 2370 records from 2008 until 2018 of wild boar in the eastern provinces of Flanders (Limburg (excluding the geographically isolated municipality of Voeren), Antwerp and Flemish Brabant), we removed 986 records which were not (yet) verified and approved by experts of Natuurpunt to increase the reliability of the observation. This resulted in 1383 wild boar occurrence recordings originating from waarnemingen.be. As a second citizen science source of wild boar occurrences, hunting records were used. For each wild boar shot a hunting record has to be entered in the data portal of the Flemish Nature and Forestry Agency (ANB). Hunters have to provide information on body weight, sex, age class, in which game management unit the wild boar has been shot, etc. Since 2016 hunters have the possibility to enter the exact coordinates when they register hunting records. Although a large part of hunters do not record the exact coordinates, a dataset of 1510 records having exact geographic coordinates of the place where wild boar were shot was available for the period from 2016 until 2018 (Figure 6.1).

Environmental variables

Based on the current knowledge about wild boar habitat use and their spatial behaviour in Europe, a set of nine land-use variables were selected (Table 6.1). The required information was retrieved from the land use map of Flanders
NARA level 2 (Poelmans and Van Daele 2014), the yearly agricultural crop maps (EPR, from 2008 until 2017 (EPR of 2018 was not yet available)) and the map of stagnant water surfaces in Flanders (including pools, puddles, ponds, fens etc. thus not including rivers, streams and canals (Packet et al. 2018)). The percentage of deciduous forest, coniferous forest, scrub and other low natural cover (natural grasslands, heathlands, wetlands, reeds), urbanized area and stagnant water in each 1 km² UTM-grid cell was calculated. Yearly agricultural crop maps were used to calculate the mean percentage of maize, the mean percentage of grasslands and the mean percentage of other crops from 2008 until 2017 in each grid cell. To assess the importance of habitat diversity, we calculated the Shannon index of habitat diversity (including deciduous forest, coniferous forest, maize, scrub and other low natural cover) per grid cell (Cornelis and Hermy 2004). All metrics were calculated in ArcMap (ESRI, 2019).
Table 6.1: Nine habitat and land-use variables calculated per 1 km² UTM-grid cell within the training area (minimum convex polygon encompassing the data records, Figure 6.1) and in Flanders with the mean percentages of these variables in a UTM-grid cell and standard error (SE) of each variable (except for Shannon habitat diversity index which represents the mean index). Studies which report the importance of these variables in spatial use of wild boar are mentioned.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean ± SE training area</th>
<th>Mean ± SE Flanders</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coniferous forest</td>
<td>8.71 % ± 0.22</td>
<td>3.96 % ± 0.094</td>
<td>Bosch et al. (2014a), Bosch et al. (2014b), Keuling et al. (2009), Morelle et al. (2016), Thurfjell et al. (2009)</td>
</tr>
<tr>
<td>Deciduous forest</td>
<td>8.38 % ± 0.15</td>
<td>5.84 % ± 0.085</td>
<td>Keuling et al. (2009), Morelle et al. (2016), Thurfjell et al. (2009)</td>
</tr>
<tr>
<td>Scrub and other low natural cover</td>
<td>2.84 % ± 0.11</td>
<td>1.72 % ± 0.047</td>
<td>Alexander et al. (2016) &amp; Bosch et al. (2014a)</td>
</tr>
<tr>
<td>Grasslands</td>
<td>14.79 % ± 0.17</td>
<td>17.68 % ± 0.11</td>
<td>Keuling et al. (2009) &amp; Morelle et al. (2016)</td>
</tr>
<tr>
<td>Maize</td>
<td>11.89 % ± 0.15</td>
<td>13.12 % ± 0.092</td>
<td>Keuling et al. (2009) &amp; Morelle et al. (2016)</td>
</tr>
<tr>
<td>Other crops</td>
<td>13.70 % ± 0.24</td>
<td>16.83 % ± 0.15</td>
<td>Keuling et al. (2009) &amp; Morelle et al. (2016)</td>
</tr>
<tr>
<td>Urbanized area</td>
<td>4.9 % ± 0.11</td>
<td>5.08 % ± 0.063</td>
<td>Stillfried et al. (2017a)</td>
</tr>
<tr>
<td>Water</td>
<td>1.47 % ± 0.068</td>
<td>1.12 % ± 0.031</td>
<td>Ilse and Hellgren (1995)</td>
</tr>
<tr>
<td>Shannon-Wiener habitat diversity index</td>
<td>0.71 % ± 0.005</td>
<td>0.52 % ± 0.0032</td>
<td>/</td>
</tr>
</tbody>
</table>

As Flanders has mainly a flat or gently undulating landscape, we did not include variables to characterize topography. Moreover, as Flanders is a rather small area with limited variation in weather conditions (i.e. to have an effect on wild boar), no climatic variables were included.

The degree of multicollinearity between these nine variables was calculated by computing pairwise Pearson’s correlation coefficients (r-value) in R (R Development Core Team 2015) (R studio Team 2016). The highest correlation...
was found to be 0.5 (between scrub and other low natural cover and other crops, Supplementary materials, Appendix 6.1) which we considered not too high to interfere model construction so no variables were excluded.

**Species Distribution Model**

As we have presence-only datasets of wild boar occurrences we opted to use MaxEnt (maximum entropy modelling) (Phillips et al. 2006) to conduct a SDM analysis. MaxEnt is a popular modelling method as it is known to produce robust models and to have a high predictive performance (Elith et al. 2006). MaxEnt models compare environmental characteristics at sites where species have been recorded with those throughout the modelled region (defined as background) (Guillera-Arroita et al. 2015, Phillips et al. 2009). MaxEnt’s predictions are indices of habitat suitability (Merow et al. 2013). MaxEnt assumes all locations in the landscape to be equally likely to be sampled (Merow et al. 2013). However, due to spatially unequal sampling effort and resulting sampling bias, some environmental variables risk to be overemphasized (Kramer-Schadt et al. 2013). Both of our citizen science data sources present a different type of bias for which correction is essential.

**Sampling bias waarnemingen.be**

Given the nature of waarnemingen.be as a data portal for all species, these data gave a unique opportunity to better understand sampling bias in such data. On the basis of sightings reported, observers can be classified in groups (i.e. mammal specialists, butterfly specialist, generalists, etc.). Not all types of observers record all species they see equally. After defining what other species an observer who submits wild boar records typically also reports, we can create a wild boar observer profile and assess sampling effort of wild boar observers throughout Flanders. A consistent wild boar observer was defined as an observer who recorded wild boar on at least 5 separate days (see Supplementary materials, Appendix 6.2 for more detailed information). The sampling effort by wild boar observers is calculated as the total number of days for which mammal observations in a UTM 1x1 km grid cell are submitted (see sup Supplementary
materials, Appendix 6.2 for more detailed information). This sampling effort assessment resulted in a raster file that can be used in a MaxEnt model to select background data corrected for sampling bias: background data are selected with a higher probability in areas with higher sampling effort than in areas with lower sampling effort.

**Sampling bias hunting bag**

Wild boar hunting in Flanders is only allowed on those properties for which the hunter has the hunting rights – so called hunting grounds. Outside these hunting grounds, there is no hunting so no wild boar shot can be reported, this results in areas without sampling effort. A map of hunting grounds is therefore used to construct a raster file based on the percentage of hunted area in each 1x1 km UTM grid cell. This raster is used in MaxEnt to select background data corrected for sampling bias: background data are selected with a higher probability in grid cells with a high percentage of the area being hunted.

**Spatial thinning**

To assess if defining sampling effort using observers’ characteristics of waarnemingen.be or using hunting ground information is a good method to correct for sampling bias, we compared the results of these approaches with those of spatial thinning or filtering. This method is often used to handle sampling bias (Kramer-Schadt et al. 2013, Swinnen et al. 2017). Applying spatial thinning, environmental variables do not risk being overemphasized. This over-representation is possible when recorded species occurrences are spatially clumped resulting in spatial autocorrelation (Kramer-Schadt et al. 2013). Spatial thinning is applied by removing data from waarnemingen.be or hunting bag data closer than the minimum “nearest neighbour index distance”. The latter is the ratio between the observed distance and the expected distance, being the average distance between neighbours whenever the observation would be random distributed thus not spatially autocorrelated). Using the *spThin* R package (version 0.1.0, Aiello-Lammens et al. 2015), this nearest index distance was found to be 0.23 km for the waarnemingen.be dataset and 0.22 km for the hunting bag dataset. Applying spatial thinning resulted in a thinned dataset of
627 occurrences of waarnemingen.be and 729 occurrences based on hunting bag records. Background data selection based on this spatial thinned dataset happened at random.

**MaxEnt**

All background data were selected within the minimum convex polygon (MCP) encompassing all data. The resulting MCP used as training area encompassed 4940 km² (Table 6.1 & Figure 6.1). Different MaxEnt models were constructed (Figure 6.2). The first two MaxEnt models were based on waarnemingen.be data and background data selection for these models was done taking the sampling bias grid based on these data into account. A first model was constructed using the full dataset and validated using hunting bag dataset. A second model was constructed using a random selection of 70% of the waarnemingen.be dataset and validated using the remaining 30%. Similarly, two MaxEnt models based on hunting bag data, and its corresponding sampling bias file for background selection, were built: a first model was constructed using the full hunting bag dataset and validated using the waarnemeng.be data and a second model was constructed using a random selection of 70% of the hunting bag dataset and validated using the remaining 30%. Using this same methodology, 4 MaxEnt models were constructed based on the spatially thinned datasets of or waarnemingen.be or hunting bag. For these 4 models, background data are randomly selected (Figure 6.2).
Figure 6.2: MaxEnt models, datasets, training and validation data selection and background selection with WA being data from waarnemingen.be, HU being data from hunting bag, Fu being the full dataset of either WA or HU, Th being the thinned dataset of either WA or HU, SBF being sampling bias file for background selection and RBS being random background selection.
Before running the MaxEnt models, the optimal settings were defined using the ENMeval R package (version 0.3.0, Muscarella et al. 2014). In this ENMeval R package, different methods are provided for partitioning training data. Our goal is to conduct predictions for the whole of Flanders, thus involving model transfer across space. We therefore used the block separation method for partitioning training data as this method was found suitable for studies involving model transfer across space (Muscarella et al. 2014). The block method partitions training data according to latitude and longitude into 4 geographically separated parts. We tested a range of different settings and their combinations using random background selection (1500 points). A regularization multiplier varying from 0.5 to 4, using 0.5 step intervals (higher values result in stronger smoothing and less complex models) and feature classes varying between linear (L), quadratic (Q) and product (P) or a combination of these classes (more classes enable more flexible and complex fits to the observed data) (Muscarella et al. 2014). Model performances of all possible setting combinations were compared using the AUC-value (area under the receiver-operating characteristics curve or ROC-curve). The settings of the models with highest AUC-value were considered optimal settings and were further used in the final MaxEnt models. With optimal settings, all models (Figure 6.2) were fitted using MaxEnt version 3.3.3 in the dismo R package (version 1.1-4, Hijmans et al. 2017). Following the method of Marchi and Ducci (2018) to evaluate the robustness of each variable delivered by a model, each model was ran 50 times, each times randomly splitting the datasets into 70% training and 30% validation (for the models using this approach, models using the full dataset have the same full dataset over the 50 runs) and for each run 1500 background points were selected based on the sampling bias files or at random. Model performance was analysed using the AUC-values averaged over the 50 runs. Based on each of these models, habitat suitability projections were made for the rest of Flanders and are then averaged over the 50 runs. The correlations between different model predictions are tested using the Pearson correlation layerStat function of the raster R package (version 2.8-19, Hijmans et al. 2019). Combining all 8 MaxEnt models, the average, minimum and
maximum possible habitat suitability was calculated. The mean variable importances over all eight models are calculated.

Extrapolation outside the training range of a SDM can result in less reliable predictions (Fitzpatrick and Hargrove 2009). To get an idea of the uncertainty of extrapolation outside the training area, the extent of environmental differences between model training and projection area can be calculated using multivariate environmental similarity surface (MESS) maps. MESS-analysis measures the similarity between the dataset used to train the model and the newly projected areas on variable at a time (degree of extrapolation of univariate ranges for individual variables). However, these MESS-maps do not visualize multivariate combinations of environmental conditions which are not represented in the dataset. We therefore used the proposed method of Zurell et al. (2012) to determine environmental overlap as an extended MESS-analysis.

By determining environmental overlap, parts of the environmental range of variables in Flanders which are within the sampled, univariate range of individual variables of the training set but which represent new multivariate combinations are identified. This is done by splitting training data into a 3 bins in which each bin holds a unique combination of environmental predictor values. Bins in the predictions dataset that do not overlap with training bins are defined as novel environments in which model extrapolation occurs (Zurell et al. 2012).
Chapter 6

Results

For all models, optimal settings were found to be a regularization multiplier of 0.5 and feature classes LQP. Model performances (defined by averaged AUC-values) of all models ranged between 0.78 and 0.88 (Table 6.2). The highest AUC values were found when trained and validated set had the same source of species occurrences. Small standard deviations reflect highly robust variables used for the Maxent models.

The percentage of coniferous forest is the most important variable defining habitat suitability over all models (Table 6.2). Furthermore, deciduous forest, scrub and other low natural cover, water and maize showed intermediate importance. The remaining variables only resulted in a minor contribution to habitat suitability. With increasing percentages of both forest types and scrub and other low natural cover, habitat suitability increases (Supplementary materials, Appendix 6.3).
Table 6.2: Variable importance (%) of all variables in MaxEnt models together with AUC-values as a measure of model performances. These variable importances are averaged (including standard deviation (SD)) over the 50 model runs for each of the eight models. Variable importances are scaled to the AUC-value (the sum of all values equals the AUC-value, not 100%). The last column presents the mean over all models with the combined SD.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Waarnemingen.be</th>
<th>Hunting bag</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>Thinned</td>
</tr>
<tr>
<td>Training-validation</td>
<td>WA-HU 70-30</td>
<td>WA-HU 70-30</td>
</tr>
<tr>
<td>Model name</td>
<td>Fu WAHU Th WAHU Th WA-AWA Th WA-AWA Fu HUWA Th HUWA Th HUWA Th HUWA Th HUWA Th HUWA</td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>0.74 ± 0.01</td>
<td>0.79 ± 0.06</td>
</tr>
<tr>
<td>Variable importance (%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
High Pearson correlation coefficients between prediction maps reflect the high similarity between predicted habitat suitability based on the different models (Table 6.3, Supplementary materials, Appendix 6.4). The averaged habitat suitability over all eight models shows that highest suitable areas were found in the East of Flanders (Figure 6.3). Towards the West of Flanders, the overall habitat suitability was found to be lower (Figure 6.3). However, small patches of highly suitable habitat distributed in a matrix of less suitable habitat are found all over the region and occur in high numbers in the centre of Flanders. Based on the averaged habitat suitability over all eight models, although of the total area, 3.75% is currently occupied (based on all waarnemingen.be and hunting bag occurrences), a remaining 8.37% of suitable habitat (habitat suitability > 0.5) is not yet occupied (Table 6.4). Of the total area, suitable habitat (habitat suitability > 0.5) raises from 7.67% to 19.21% when minimum and maximum habitat suitability are compared (Supplementary materials, Appendix 6.5).

Our environmental overlap MESS-analysis showed that model projections towards the rest of Flanders involved predictions towards novel environments in the West and Southeast of Flanders (Supplementary materials, Appendix 6.6) indicating extrapolation of our models in these areas.

<table>
<thead>
<tr>
<th></th>
<th>FuHUHU</th>
<th>FuHUWA</th>
<th>FuWAHU</th>
<th>FuWAWA</th>
<th>ThHUHU</th>
<th>ThHUWA</th>
<th>ThWAHU</th>
<th>ThWAWA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FuHUHU</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FuHUWA</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FuWAHU</td>
<td>0.98</td>
<td>0.98</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FuWAWA</td>
<td>0.98</td>
<td>0.98</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ThHUHU</td>
<td>1</td>
<td>1</td>
<td>0.98</td>
<td>0.98</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ThHUWA</td>
<td>0.80</td>
<td>0.80</td>
<td>0.76</td>
<td>0.74</td>
<td>0.80</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ThWAHU</td>
<td>0.84</td>
<td>0.84</td>
<td>0.80</td>
<td>0.78</td>
<td>0.84</td>
<td>0.98</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>ThWAWA</td>
<td>0.83</td>
<td>0.83</td>
<td>0.80</td>
<td>0.78</td>
<td>0.83</td>
<td>0.98</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 6.3: MaxEnt prediction of habitat suitability for wild boar in Flanders averaged over all eight models. Red colours indicate high habitat suitability; green colours indicate low habitat suitability. Dots are waarnemingen.be and hunting bag occurrences. Model predictions of all separate models can be found in Supplementary materials, Appendix 6.4.

Table 6.4: Occupied versus unoccupied suitable habitat (defined as habitat suitability > 0.5) in Flanders based on the averaged habitat suitability over all eight models. Percentages are defined as the percentage of grid cells being occupied or unoccupied by all waarnemingen.be and hunting bag occurrences.

<table>
<thead>
<tr>
<th>Percentage in Flanders</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Suitable area</td>
<td>12.14 %</td>
</tr>
<tr>
<td>Occupied area</td>
<td>3.75 %</td>
</tr>
<tr>
<td>Unoccupied suitable area</td>
<td>8.37 %</td>
</tr>
</tbody>
</table>
Discussion

By relating wild boar occurrences to environmental variables, we obtained a better understanding of factors influencing habitat suitability for wild boar in the highly anthropogenic landscape of Flanders. Based on our model projections, remaining suitable habitat was mainly found in the East of Flanders adjacent to the current distribution area and in some areas in the centre of Flanders. Towards the West of the region only a limited amount of suitable habitat was found.

Forest played an important role in defining habitat suitability. This is in line with previous studies who found that forest plays a key role in landscape use and range expansion (Alexander et al. 2016, Morelle et al. 2016, Rutten et al. 2019b). Coniferous forest was found more important than deciduous forest in defining habitat suitability. While some studies did not consider separate forest types having a different contribution (i.e. they did not differentiate between different forest types) (Alexander et al. 2016, Bosch et al. 2014b), Thurfjell et al. (2009) in Sweden and Fonseca (2008) in Poland reported deciduous forest more preferred compared to coniferous forest. A potential reason mentioned by Thurfjell et al. (2009) and Fonseca (2008) is the high presence of food resources in deciduous forest through mast. These findings are in contrast to ours. A potential explanation may be that intensive agricultural systems like the one in Flanders (also found in the Netherlands, Denmark, parts of Germany, France etc.) provide a surfeit of food outside forests leading wild boar populations to be less dependent on mast in deciduous forest. Although scrub and low cover is generally considered not improving habitat suitability due to limited shelter opportunities (Alexander et al. 2016, Bosch et al. 2014a), a considerable contribution of this landscape type to habitat suitability was found. A possible explanation may be that this is an adaptation due to limited and fragmented forest availability: while scrub and low cover areas are often considered as marginal habitats, we think that wild boar can find sufficient shelter in this vegetation allowing them to move through a highly anthropogenic landscape. Moreover, providing shelter in forests might be
determined specifically by undergrowth in forests: open forests provide less shelter than forests with scrub-like undergrowth. However, as there were no available maps on undergrowth cover in forests, this could not be assessed. The relationship of both the percentage of maize and water with habitat suitability was found not to be uniform between models based on waarnemingen.be data and hunting bag data. However, high correlations between habitat suitability projections of all eight models were found illustrating the overall consistency of our results.

The Flemish landscape is characterized by a much higher anthropogenic pressure than study areas of previous conducted studies which allowed us to assess if there are limits to the flexibility of wild boar towards habitat use. Urban cover did not have a large negative impact on habitat suitability. It has been shown that wild boar show substantial behavioural plasticity to adjust to human-dominated environments i.e. landscape of fear (Stillfried et al. 2017a): wild boar tolerate human presence by modulating their risk perception indicated by lower flee distances of urban boars versus rural boars, adjust their spatial use (use of recreational forest with high human presences) and even use human-associated habitat classes by modulating the perception of harmless anthropogenic risk. Habitat diversity was found not important confirming adaptive and highly flexible habitat utilization by wild boar in response to anthropogenic disturbances (Calenge et al. 2002, Keuling et al. 2008c, Maillard and Fournier 2014). These results show that we have not yet reached the limits of wild boar being able to use human-dominated habitat types or habitat like scrub and other low natural cover which are generally not considered to improve habitat suitability when adjusting to anthropogenic landscapes. We therefore acknowledge that given the high behavioural plasticity of wild boar in adjusting towards human-dominated areas, wild boar might be able to further adjust to less suitable areas, to changing environments and changing climate (Lowry et al. 2013). MESS maps showed extrapolation to novel environments in the East of Flanders where the landscape is dominated by agriculture and there are limited areas covered by forests. If wild boar can further adjust to
these open novel environments, it is possible that areas which are currently considered not suitable may still become colonised.

Current model projections showed the existence of yet still unoccupied highly suitable habitat, although generally scattered throughout the landscape, in the East and parts of the centre of Flanders. On the other hand, the currently unoccupied habitats in the West are less suitable for wild boar. Although population numbers and distribution range expanded fast since their recolonization in 2006 (Scheppers et al. 2014), based on the models one would expect limited future expansion of the current wild boar distribution range in Flanders as only 12.14% of the Flemish landscape exists out of suitable habitat. However, since currently only 3.75% of Flanders is occupied, only one third of the suitable habitat is currently used so the current expansion range can still triple in area.

To study large-scale patterns, a large amount of data needs to be collected (Bonney et al. 2009). Using citizen science of waarnemingen.be we did not only have a large amount of wild boar occurrence data, but also information on search effort derived from other wildlife observations. This allowed defining sampling effort in a new way to create bias files for background selection. However, we want to acknowledge that differences in detection probabilities of different kind of occurrence data (sightings, footprints etc.) were not quantified and thus not assessed if this affected modelling results as this was outside the scope of this study. Furthermore, by using information on hunting grounds we could correct for sampling bias in the models based on citizen science data originating from hunting bag information. By comparing these new methods to the already used method of spatial thinning, we found very similar outcomes with comparable model performances. Analysing observers’ characteristics based on their observation recorded in waarnemingen.be has previously been found useful when calculating search effort corrected population trends in butterflies and birds (Herremans 2010). Although spatial thinning involves less effort than defining sampling bias, spatial thinning reduces the amount of data which can be problematic for small datasets for example of species with a low
detection probability or datasets of rare species. In these cases, creating bias files using our presented method can offer a solution. Moreover, as hunting bag data was found to be an important source for large-scale SDM for wild boar in Europe, given the fact that they are the most available and standardized source of wild boar occurrences throughout Europe (ENETwild consortium et al. 2019), correcting sampling bias based on using hunting ground information can be useful and easily applicable, also on larger scales.

Management implications

Wild boar is expected to expand its distribution range in Flanders. As HWI in an anthropogenic landscape such as Flanders strongly affects stakeholder acceptance of wild boar, being able to conduct risk assessments linked to the future wild boar expansion is essential to assess the future possible evolution of HWI. A risk assessment allows prioritizing management actions in areas where wild boar is expected due to high habitat suitability. Moreover, crop damage is one of the HWI raising most concerns because of the high economic impact (Carnis and Facchini 2012, Schley et al. 2008). Recent research of Rutten et al. b (not in press) assessed landscape factors influencing crop damage probability in Flanders. However, since these damage probability predictions did not yet include future distribution of wild boar, creating this SDM of wild boar in Flanders is an essential step to predict the geographic distribution and extent of damage risks linked to further wild boar expansion. Using the combined information on habitat suitability and damage probability within areas occupied by wild boar, implementing preventive measures where the highest damage risks are localised, can allow to decrease crop damage and thus increasing stakeholder acceptance. Furthermore, recent outbreaks of African Swine Fever (ASF) in several eastern European countries including an outbreak in Wallonia in September 2018 raises concerns due to major economic impacts (Costard et al. 2009, Lange et al. 2018). The gained understanding on the future expected distribution of wild boar in Flanders will be essential towards a risk assessment with potential future ASF outbreaks.
Acknowledgements

A. Rutten holds a PhD grant (IWT.141377) from the Agency Flanders Innovation & Entrepreneurship (VLAIO), Flanders, Belgium. We thank all citizen scientists for their recordings. Furthermore, we thank all the species experts for validating observations in the data portal ofwaarnemingen.be We thank hunters for recording specific coordinates of hunted wild boar. Finally we want to thank dr. Dirk Maes and Emma Cartuyvels for their helpful advice on the manuscript and analysis.
Chapter 7

Discussion
Although wild boar have been part of Flemish wildlife for thousands of years, their re-emergence in 2006 after local extinction for half a century causes concerns. Economic growth, urbanization and the intensification of agricultural practices resulted in the fragmentation of the Flemish landscape, currently one of the most densely human populated areas of Europe. Because of this anthropogenic landscape, wild boar often comes into contact with human activities. Both actual Human-Wildlife Impacts (HWI) and fear for more HWI currently limit stakeholder acceptance of wild boar as part of the Flemish wildlife.

This PhD thesis addressed several key aspects needed to conduct a risk assessment related to the current and possible future distribution of wild boar in Flanders. The main goal was to provide a scientific basis to guide future management planning to decrease Human-Wildlife Impacts. This final chapter will connect the findings of the conducted research and discuss their contribution towards a risk assessment. On the way to our goals, we met methodological challenges, had to balance available with desired data and several additional research questions and interests arose. Coming towards the end of this thesis, we are convinced that our research provides an essential contribution towards a risk assessment, guiding management planning and a more efficient implementation of preventive measures.

**Crop damage**

Our drone-based damage assessments using geographic object-based image assessment (GEOBIA) resulted in a highly accurate method to define the damaged area in a particular field (Chapter 3). By comparing this method to ground visits like done in Wallonia (J. Widar, personal communication), we showed that our method is more cost-efficient to apply. Furthermore, we showed that our method is more time-efficient than a detailed ground assessment. This method can therefore certainly be implemented to monitor crop damage in Flanders and elsewhere. However, airspace restrictions do
prohibit drone use in certain areas at the moment or require special permits to allow the use of drones.

Without a regular check within the field, crop damage is sometimes only observed during harvest, given that wild boar often only damage maize in the centre of a field. Therefore, monitoring a large area for damaged fields by the use of satellite imagery might give a clear view on the actual number of damaged fields. Satellite images generally have a low resolution (for example free SENTINAL-2 satellite imagery has a resolution of 10 m (Copernicus, ESA)). By characterizing damaged and undamaged fields on low resolution using the GEOBIA methodology, a monitoring method could be developed. In large fields of several hectares, a rough estimation of the damaged area within a field might be feasible. However, as Flemish agricultural fields are generally small (on average 1.4 ± 1.8 ha), a detailed assessment of the damaged area within a field is not achievable due to the low resolution. Moreover, due to seasonality of agricultural crops and damage which can occur just before harvest (Schley et al. 2008), persistent cloud cover can prevent cloud-free satellite images for a certain time period and location (Koh and Wich 2012). These factors prevent an efficient use of satellite images to assess crop damage within fields with high resolution. The combination of satellite-based monitoring and detailed drone-based damage assessments could give an overview on the actual extent of crop damage in an area.

Chapter 2 described the results of an online survey questioning farmers if they already experienced damage by wild boar and a follow-up non-respondent check to assess potential respondent bias. Our results showed that the online survey gave a general overestimation of crop damage probability, while the non-respondent check demonstrated a general underestimation. Although we could document the current spatial distribution of crop damage in the eastern part of Flanders, the revealed respondent bias made it impossible to draw conclusions on the factors explaining the spatial distribution of crop damage. We therefore focused on this question using detailed information about the characteristics of a field as well as about the landscape surrounding the field.
in Chapter 4. As far as concerns the probability of damage to maize fields, the landscape characteristics were the dominant factors and field-specific factors only played a minor role. In contrast, damage probability in grasslands was mainly influenced by field-specific factors and the landscape surrounding a field only played a minor role. The latter resulted in landscape predictions which were less useful. However, the large influence of field-specific factors for grasslands does allow farmers to adjust agricultural practices like fertilizing schemes to maximally prevent crop damage. For maize fields on the other hand, adjusting these practices has less effect which makes the use of preventive measures like (electric) fences more profitable. However, in both cases, costs are involved to take preventive measures or economic losses can be expected by adjusting fertilizing which might causes lower yields (Di Paolo and Rinaldi 2008, Goetz 1969, Schläpfer et al. 2002). The increased knowledge based on our research, allows to better evaluate options to minimise damage case by case.

High hunting pressure in the surroundings of maize fields reduced the damage probability. This effect of reduced damage with increased hunting activity was also reported earlier by Geisser and Reyer (2004). In Flanders hunting is only allowed on hunting grounds where a hunter has the hunting right. In nature reserves, hunting can be prohibited as it can cause disturbance to certain species, moreover landowners not always allow hunting. As a consequence, not all maize fields are part of hunting grounds or have hunting activities in the surroundings, so only a limited number of maize fields can thus profit from this damage reducing effect of hunting. Moreover, since different hunting methods have been found to affect spatial patterns differently, ranging from decreased home ranges to increased fleeing distances (Keuling et al. 2008c), the effect of hunting pressure and different hunting methods on crop damage patterns might be more complex therefore requiring a more detailed investigation.
Fragmentation

That small forest fragments can provide sufficient shelter for wild boar to move through a landscape matrix to reach agricultural fields was found by Herrero et al. (2006) and Thurfjell et al. (2009). However these studies were conducted in landscapes of southern Sweden and Spain which have a low anthropogenic pressure given the low amount of urban and housing areas and low road density. We therefore assessed the role of habitat fragmentation on crop damage probability in the severely fragmented landscape of Flanders (Chapter 4). **We did not find an effect of habitat fragmentation on damage probability**, reflecting that landscape fragmentation within the range we observed in Flanders does not influence damage probability.

The role of landscape connectivity in fragmented landscapes such as Flanders was also assessed based on a landscape genetic analysis (Chapter 5). By focusing on forest fragmentation and road density together with different types of land cover, we found an **important effect of the percentage forest on landscape connectivity**: increasing forest percentages decreased genetic distances. However, the degree of fragmentation of forest complexes did not affect genetic distances at all, nor was there an effect of road density. Large motorways have previously been found not to pose a barrier to wild boar based on genetic structure of wild boar in Wallonia (Frantz et al. 2012). Additionally, a more recent study in Wallonia could not find an effect on genetic diversity of motorways acting as a barrier to dispersal although they did find significant barrier effects using capture-mark-recapture (Dellicour et al. 2019). We assume that roads do not act as absolute barriers allowing occasional crossings. Therefore barrier effects are detected immediately using capture-mark-recapture but are not detected using genetic approaches. Our findings confirm the substantial behavioural plasticity regarding adjustment to human dominated environments (Stillfried et al. 2017a). **We found no evidence that wild boar recolonization or damage patterns are influenced by landscape fragmentation at the scale that we conducted our study.**
It must be stressed, however, that these results are valid within the relatively narrow range of high degrees of fragmentation in the Flemish landscape; over a larger range of landscapes, for example comparing the vast and relatively undisturbed areas in Scandinavia with the situation in Flanders, the relation between fragmentation and damage probability and wild boar distribution may be different.

Distribution

In line with geographic patterns of the hunting bag, our genetic analyses confirmed that recolonization started with two distinct genetic clusters in 2007: a more eastern and a more western population. Starting from these two founder populations, distribution ranges extended in the next years. With the available dataset, we could not identify the source of these two Flemish founding populations. The origin of these Flemish founding populations is a much debated subject in Flanders as natural migration as source for these two founder populations can be excluded as these populations were geographically not connected to neighbouring populations thus implying (illegal) reintroductions or escapes from enclosures. A more extensive analysis including samples from other regions might give more information on their genetic origin as most wild boar populations in Europe are genetically differentiated in subpopulations (Scandura et al. 2011a). However, potential translocations for hunting purposes, release of captive animals and potential inbreeding with domestic pigs can make the search towards their exact origin complex (Scandura et al. 2011a, Scandura et al. 2011b). It could be questioned if it would still be useful to conduct a more extensive analysis towards their genetic origin as since 2015 a third genetic cluster emerged in the North of Flanders. This cluster is connected to Dutch/German populations suggesting natural migration so it was only a matter of time until neighbouring population expansions reached Flanders resulting in the same management situation and impacts. Moreover, even when the identification of the genetic origin is possible, this does not necessarily give an answer on who or how the wild boar were reintroduced in Flanders. However,
this matter has a large societal importance due to risks induced by translocations such as introducing disease reservoirs.

The management and control of wild boar strongly benefits from detailed spatial information on their distribution (ENETwild consortium et al. 2019, Pittiglio et al. 2018). Because landscape fragmentation does not seem to affect wild boar distribution, future range expansion can be expected as long as there is still suitable habitat available. Our habitat suitability model using species distribution modelling (Chapter 6) showed that this still is the case. Other European studies in north-eastern Germany, southern Sweden, Wallonia and Bulgaria reported the positive effect on wild boar habitat suitability of forest cover (often called the main driver of wild boar expansions) and agricultural areas which can offer both food and shelter opportunities (Bosch et al. 2014a, Bosch et al. 2014b, Keuling et al. 2009, Morelle et al. 2016, Thurfjell et al. 2009). However, the landscapes in these studies were characterised by higher proportions of forest, less fragmentation and lower anthropogenic pressures. We assessed if habitat requirements in the very divergent Flemish region with high anthropogenic pressure, differs from these less fragmented landscapes to determine the flexibility of wild boar in adjusting towards human-dominated landscapes. It turned out that like elsewhere, coniferous- and deciduous forest and maize availability were among the most important environmental factors defining habitat suitability for wild boar in Flanders. So even in such an anthropogenic landscape, the same factors were found as in very divergent landscapes. These environmental elements, being those that provide shelter opportunities like forests and seasonal agricultural crops, remain essential in defining habitat suitability in a highly anthropogenic landscape.

Scrub and low natural cover are generally not considered to advance habitat suitability (Alexander et al. 2016, Bosch et al. 2014a). However, we found a positive effect of scrub and low natural cover affecting habitat suitability which can be the result of a behavioural adaptation by using scrub and low natural cover due to limited and fragmented forest availability: wild boar can find sufficient shelter in this vegetation allowing them to more through a highly
anthropogenic landscape. Habitat use is a trade-off between foraging and the risk perception of (human) disturbance or predation risk: animals have the ability to learn and respond to varying levels of predation risk or disturbances by adjusting their behaviour (Laundré et al. 2010). Highly anthropogenic areas support various food sources such as a surfeit of agriculture crop due to intense agriculture. As forests providing shelter are limited in Flanders, wild boar use scrubs and low natural cover to move through the landscape of fear while foraging. This adjustment of the perception of the landscape of fear was previously also shown by Stillfried et al. (2017a) in Berlin. This adaptation might thus be the result of wild boars’ behavioural flexibility in adjusting towards anthropogenic landscapes.

The combination of favourable environmental factors indicates high habitat suitability in the East and parts of the centre of Flanders and limited habitat suitability in the West. This raises the expectation that although wild boar was able to expand fast since their recolonization in 2006, their future expansion will not cover the whole of Flanders. However it might still be that we have not yet reached the limits of wild boar behavioural adaptability to adjust to anthropogenic landscapes. In this case, we cannot exclude the possibility that wild boar might be able to colonize habitats which are currently regarded to be less suitable like the West of Flanders. If future distribution of wild boar results in less suitable habitats being colonised, the current model should be re-evaluated and could result in a better insight into these mechanisms. Moreover, although we see a general higher suitability in the East and centre of Flanders, suitable areas are scattered throughout the landscape. This can potentially prevent further distribution from the current distribution range to suitable areas which are surrounded by unsuitable habitat. However, the potential ability to adjust to less suitable habitats, even if this would be of a temporarily nature, can allow wild boar movements through unsuitable areas to reach scattered areas with high suitability which facilitates future distribution.
Towards a risk assessment

Based on our assessment of drivers influencing the risk on crop damage, crop damage risk maps were made in Chapter 4. However, these maps do not take into account the actual and potential future distribution range of wild boar. The habitat suitability analysis of Chapter 6 gives us a better understanding of the potential future distribution range. Combining the information of both chapters therefore allows us to evaluate expected damage risk under the potential future distribution of wild boar. For grasslands such analysis is less useful due the high influence of field specific variables. However, for maize fields in which damage probability is mostly defined by landscape specific variables, such analysis can provide an accurate overview on expected damage probability for maize fields taking into account habitat suitability throughout Flanders. We calculated the expected damage probability value of each field for the case maize is grown in that specific field and taking into account habitat suitability of the area around the field. This was done by multiplying the damage probability in maize fields (from Figure 4.2 in chapter 4 ranging from 0 to 100%) and the value of habitat suitability (from Figure 6.3 in chapter 6, ranging from 0 to 100) of the grid cell in which the field is located (Figure 7.1).
Figure 7.1: Expected maize damage probability in Flanders as a result of the multiplying between habitat suitability (chapter 6) and maize damage probability based on landscape factors (chapter 4).

By combining actual habitat suitability with each fields’ damage probability, we generally find a lower damage probability than found in Chapter 4 (Figure 7.1). Highest crop damage probabilities are found in the East and North of Flanders and a general low damage risk is found in the West and South. Based on this expected damage probability, we can distinguish 4 groups of maize fields: fields with high damage probability (>50 %) in a highly suitable area (>50), fields with low damage probability (<50 %) in a highly suitable area (>50), fields with high damage probability (>50 %) in an unsuitable area (habitat suitability <50) and fields with a low damage probability (<50 %) in an unsuitable area (habitat suitability <50). By conducting a Principal Component Analysis (PCA) including the most important variables from both the habitat suitability model and the damage probability model, we can visualize how these 4 different kinds of fields can be distinguished (Figure 7.2).
Figure 7.2: Principal Component Analysis (PCA) of expected maize damage probability in relation to the most important variables defining habitat suitability (water availability: Hab_Water, scrub and other low cover: Hab_Low_Cover, maize availability: Hab_Maize, coniferous forest: Hab_Coni_Forest and deciduous forest: Hab_Deci_Forest) and maize damage probability (percentage of urban area: Dam_Urban, percentage of grasslands: Dam_Grasslands, percentage of maize fields: Dam_Maize, percentage of forest: Dam_Forest). Fields with high damage probability (Dhigh) in a high suitable area (Hhigh), fields with high damage probability (Dhigh) in a low suitable area (Hlow), fields with low damage probability (Dlow) in a high suitable area (Hhigh) and fields with low damage probability (Dlow) in a low suitable area (Hlow) are distinguished. The total percentage of these field categories in the whole of Belgium was calculated.

Forest was found important both to define habitat suitability (chapter 6) and to define crop damage probability (chapter 4); our PCA shows that fields with high damage probability in highly suitable areas are located in the most forested areas. Low damage probability fields in unsuitable area are characterized by
less forested areas. In suitable areas low damage probability fields are distinguished from high damage probability fields by lower urbanization and a higher percentage of maize in the surroundings. In unsuitable area low damage probability fields are distinguished from high damage probability fields by lower urbanization, a higher percentage of maize in the surroundings and a lower percentage of forest. Most Flemish fields are characterised by a low damage probability (66.5% in unsuitable and 1% in suitable area), a small percentage of fields have a high damage probability in suitable area (3.5%). 29% of fields have a high damage probability in low suitable area. Based on the current habitat suitability model, no damage is expected in these areas. However, as discussed, it might be possible that wild boar are able to adjust to less suitable habitats as we have not yet reached the limits of wild boar behavioural plasticity in adjusting towards anthropogenic landscapes. If it would indeed be the case that wild boar will further adjust, additionally 29% of fields in Flanders become sensitive to crop damage by wild boar.

This result in important information which can now be taken into account in a risk assessment: areas in which management actions should be taken to prevent crop damage and decrease Human-Wildlife Impacts are now clear.

Furthermore, by comparing the expected damage probability from chapter 6 (Figure 6.3) with minimum and maximum habitat suitability as a result of combining all 8 models in chapter 6 (Supplementary materials, Appendix 6.5) using a density plot, the distribution of fields with a higher expected damage probability of 50% is visualised (Figure 7.3). This density plot shows that based on minimum habitat suitability, a lower part of the fields have an expected damage probability higher than 70% compared to maximum habitat suitability where most fields have an expected damage probability higher than 70%. Overall, the total percentage of fields with an expected damage probability higher than 50% rises from 0.8% for minimum habitat suitability, to 1.4% for the modelled habitat suitability from chapter 6 to 2.4% for maximum habitat suitability.
Figure 7.3: Density plot of expected damage probability (only fields with expected damage probabilities higher than 50% are visualised). The comparison is made for the modelled habitat suitability of the FUWAWA model in chapter 6 (Modelled HS), the maximum habitat suitability (Max. HS) and minimum habitat suitability (Min. HS). The total percentage of fields with a expected damage probability higher than 50% is shown between brackets.

Our models and their projections come with uncertainties due to variables which are not included in our models (such as soil characteristics in Chapter 4), model extrapolation, imperfect model accuracies, etc. These uncertainties affect our risk projections. However, since the Flemish context is quite divergent from many regions, projecting models based on research conducted in strongly different areas, would probably induce even to a larger extent uncertainty. Our models gave valuable insights in drivers of crop damage patterns and habitat suitability in the anthropogenic landscape of Flanders. Overall, this PhD research gave an essential scientific and objective basis in order to conduct a risk assessment.
Future research

Our research opens possibilities for a more extensive assessment of certain research aspects such as financial aspects affecting economic impact of crop damage and potential density-dependent effects on future distribution and crop damage probability.

Currently, we were only able to develop our drone-based method for the two main crops suffering wild boar damage in Flanders: maize and grasslands. Although a variety of other crops are also damaged by wild boar e.g. barley, oats, triticale, wheat, potatoes, etc. (Herrero et al. 2006, Schley et al. 2008), our sample did not include sufficient damage cases to develop the method for other crops. However, as long as the subject of interest, which in our case is crop damage, is visually distinguishable (Addink et al. 2012, Blaschke 2010) in aerial photographs, we are convinced that our method can be applied to other crops. As the workflow of our method is known, a sufficient training on orthophotos of other crops will allow the development of models with comparable accuracies as for maize fields and grasslands. Similarly as done for maize and grassland, analysing the factors influencing the crop damage probability of other crops would subsequently be possible. Finally, a more extensive assessment on crop damage probability by wild boar for the current and future distribution of wild boar would then be possible.

We have shown that even in the severely fragmented landscape of Flanders wild boar were able to thrive by adjusting their habitat use to this anthropogenic landscape. Given that we have not found limits of wild boars’ behavioural plasticity in adjusting towards anthropogenic pressures, we acknowledged that wild boar might be able to settle in less suitable habitats which were assessed in Chapter 6. It might be that drivers of habitat selection or crop damage patterns become divergent between a recent colonizing population and a settled population due to evolution, adaptation, limited space, density dependent effects, etc. Assessing potential changes in habitat preference with future distribution and the comparison of these results with the current habitat
suitability characteristics (Chapter 6) might increase understanding of possible limits towards adjustment to anthropogenic pressure.

Our assessment of factors that influence crop damage probability (Chapter 4) took into account both landscape- as well as field-specific factors but as we could not include variations in wild boar densities within its range in Flanders, we were not able to analyse the crop damage patterns in relation to population densities. We assume that it might be possible that next to landscape- and field-specific factors, also wild boar densities play a role in determining damage risk. In this PhD, we did not attempt to determine population densities. Various methods have been suggested to estimate population densities of wild boar (Acevedo et al. 2007, Focardi et al. 2001, Focardi et al. 2002, Massei et al. 1997b). Each of them has particular challenges which makes it difficult to assess densities on a larger scale and often only allows to estimate indices of population changes (Engeman et al. 2013, Scheppers and Casaer 2012). Addressing the disadvantages and challenges of such methods will hopefully allow to define wild boar densities in Flanders.

Since 2014, outbreaks of African Swine Fever (ASF) in several eastern European countries causes serious concerns due to high mortality rates and major economic impacts when domestic pig farms are infected with ASF (Costard et al. 2009, Lange et al. 2018). With recent outbreaks of ASF in Wallonia in September 2018 (Portail de la Wallonie, 2019), concerns and the need to take preventive measures arise in Flanders. Research towards ASF should include habitat suitability assessments as done by Lange et al. (2018) in the Baltics and Poland which resulted in the assessment of physical barriers, the role of human-mediated translocation and most effective preventive- and control measures to prevent the spread of ASF maximally. Furthermore, simulations of epidemiological and economic effects on domestic pig industry can add in a risk assessment of ASF outbreaks (Halasa et al. 2016).
Conclusion

In conclusion, this PhD provided a sound scientific basis to conduct a risk assessment of wild boar impacts with their potential future expansion in Flanders. Wild boar distribution does not seem to be affected by habitat fragmentation, thus future expansion of the Flemish wild boar population can be expected to remaining areas with suitable habitat in the severely fragmented landscape of Flanders. Moreover, wild boar behavioural plasticity in adjusting towards anthropogenic pressures may not yet have reached its limits which could mean that wild boar may also colonise new types of habitat.

A drone-based method can be used in a time- and cost-efficient way to assess with high resolution and accuracy the damaged area within a field. This could allow follow-up of damage over time or in space, or support compensation schemes.

While crop damage probability of maize fields is mainly influenced by the landscape surrounding a field, damage risk in grassland depends mostly field-specific factors. This means that for maize fields, the landscape may indicate whether preventive should be undertaken, while for grassland local agrotechnical adjustments could be considered.

We are convinced that the results obtained in this study are not only of use in the Flemish situation but can also be applied to similar situations abroad where wild boar presence is challenging due to fragmented landscapes, close contact with human activities, economic losses due to crop damage and population increases.
Dankwoord
Dankwoord
Een van de leukste delen van een doctoraat afwerken is toch wel een dankwoord te mogen schrijven. Bijna 5 jaar lang heb ik met enorm veel motivatie en passie aan mijn doctoraat gewerkt maar de realisatie van dit proefschrift was onmogelijk geweest zonder een heel aantal mensen. Daarom zou ik hier de tijd willen nemen om een gepast bedankje neer te schrijven.

Allereerst wil ik mijn promotor Herwig bedanken. Jouw gepassioneerde uitleg op de infodag voor Biologie toen ik als 18-jarge aan het uitzoeken was wat ik met mijn leven wou doen heeft mij doen beslissen om naar de Universiteit Antwerpen te komen. Tijdens mijn bachelor opleiding heeft de fantastische Tanzania-stage me doen beseffen dat onderzoek wel iets voor mij was en de kans om mijn masterthesis over beverdammen bij jou te kunnen doen heeft mij helemaal verkocht om verder onderzoek te voeren naar grote zoogdieren en Human-Wildlife Conflicts. Ik heb dan ook geen enkel moment spijt gehad om een doctoraat te starten onder jouw begeleiding. Mijn doctoraat is ongetwijfeld een succes geworden door jouw ongelofelijke wetenschappelijke inzichten, de vele overlegmomentjes en jouw vertrouwen in mij waarvoor dank! Bedankt om mee te gaan in mijn ambitieuze ideeën zoals een drone aanschaffen, bedankt ook om bij elke crash weer financieel bij te springen om de drone terug de lucht in te krijgen en bedankt om mij telkens te pushen mijn papers en mijn proefschrift naar een hoger niveau te tillen!

Jim, zonder jou als copromotor was mijn doctoraat onmogelijk geweest! Vanaf dat ik bij jou kwam informeren over doctoraatsmogelijkheden met everzwijnen in samenwerking met het INBO was ik door jouw enthousiasme ervan overtuigd dat we een nuttig en super interessant onderzoeksproject konden ontwikkelen. Jij hebt er voor gezorgd dat ik mij ook op het INBO thuis kon voelen als leer- en werkomgeving! Bedankt om altijd voor mij klaar te staan met mijn onuitputtelijke vragen, mijn twijfels en onzekerheden en vooral heel erg bedankt om niet gek te worden van mijn koppiheid en gezaag om weer iets zo snel mogelijk na te lezen! Jouw enthousiasme en enorm doorzettingsvermogen hebben mij als doctoraatsstudent zeker in een luxepositie gezet wetende dat je altijd je gsm opneemt en even hard als ikzelf blijft.
Dankwoord

nadenken als we vast zaten met iets van mijn doctoraat. Verder kan ik je niet genoeg bedanken voor jouw vertrouwen in mijn capaciteiten als wetenschapper door mij mijn job bij het INBO aan te bieden en mij te steunen met verdere carrièreplannen.

Verder wil ik mijn vrienden en collega’s op de unif bedanken. Eerst en vooral Bram, bedankt voor de vele welgekomen pauzes, de babbeltjes, klimavondjes en fietstochtjes! Jouw humor heeft veel van mijn stressy dagen kunnen opfleuren. Bedankt ook om mij niet te hard uit te lachen met mijn onhandige toeren en valpartijtjes! Jij bent een toppertje! Joachim, mijn bureauvaartje en reisgenootje, door samen aan ons doctoraat te starten heb ik jou echt leren bewonderen als wetenschapper, jouw filosofische levenswijsheid maken een dag goed! Luci, sharing an office with you has been a pleasure, thanks for all the help and the advice for my future career, you are an excellent postdoc scientist and you will make a great career! Myrthe, bedankt voor de vele babbeltjes, lachmomentjes en zwemmotivatie, je bent een super Limburgs steun- en toeverlaat’je! Ook Kristijnn wil ik graag bedanken voor zijn super begeleiding tijdens mijn masterthesis! Dankzij jou heb ik kunnen zien wat een doctoraat inhoudt en heb ik durven de stap te zetten om er zelf aan te beginnen. Ondertussen zitten we al aan 2 gepubliceerde papers samen en ik kijk uit naar de volgende! Stijn en in mijn eerste jaar ook Michel, bedankt voor de administratieve ondersteuning, aangezien dit nooit mijn sterkste kant is geweest en ook nooit zal zijn ben ik jullie enorm dankbaar om mijn doctoraat administratief in goede banen te leiden. Furthermore, all the rest of the lab, Omid, Laura, Constance, Frederik, Nathalie, Lisa B, Mats, Babu ram, Gerardo, Vincent, Natalie, Joris, Frank, Luc, Erik, Stefan and old lab members Benny, Sophie, Dieter, Dries, Lisa A and Michalis, thanks for making EVECO such a great lab!

Een dikke merci gaat ook uit naar mijn vrienden en collega’s van het INBO. Jan, van onze epische moddergevechten met de everzwijntjes tot vele babbeltjes, zonder jou zou het FIS-team niet hetzelfde zijn! Sander, bedankt om altijd met veel enthousiasme klaar te staan voor mijn vele R- en GIS-vraagjes,
Dankwoord

top-Limburger! Thomas, bedankt om mij mee te helpen wroeten doorheen mijn genetica-onderzoek en het gezelschap tijdens de everzwijncongressen! Frank, op tijd en stond was er dat perfect getimed mopje of gevatte opmerking van jou waarvoor dank! Fauna-babes Yasmine en Emma, bedankt voor de vrouwelijke ondersteuning in ons mannenteam, onze babbeltjes met bijhorende Cava maakt een werkdag compleet! Ook de andere teamleden Kristof, Axel, Tim, Joris, Koen, Filip en Jan maken van FIS een super team waarbij ik mij vereerd voel om er deel van te mogen zijn!

Ook wil ik alle anderen bedanken die mij ergens op de weg van mijn doctoraat hebben geholpen, mij hebben bijgestaan via technische ondersteuning, geduld gehad om mij uitgerek togeven of om een bijdrage te leveren aan een van mijn publicaties: Diederik Strubbe, Jeroen Vanden Borre, Thierry Onkelinx, Lieven De Smet, Nele Witters, Marjolein Vogels, Elisabeth Addink, Karen Cox, Dirk Maes, Marc Herremans, Pieter Vandenbroucke, Stien Herremans en Jolien Wevers - hartelijk bedankt!

Enorme appreciatie gaat ook uit naar alle landbouwers en jagers die mij tijdens mijn doctoraat hebben opgebeld om landbouwschade te melden. Zonder jullie medewerking was mijn onderzoek onmogelijk geweest! Van sommigen onder jullie kreeg ik jaarlijks terugkerende telefoontjes, en hoewel dit betekende dat de everzwijnen weer op bezoek waren geweest ben ik blij dat ik jullie daardoor goed heb leren kennen! Ik hoop dat ik, door een luisterend oor te kunnen bieden en jullie bij te staan met raad, jullie een beetje heb kunnen ondersteunen.

Daarnaast wil ik ook de mensen van de Boerenbond, de Hubertus vereniging (HVV) en het Agentschap voor Natuur en Bos (ANB) bedanken voor de fijne samenwerking doorheen de jaren. Bedankt voor het aanleveren van data en om mij uit te nodigen op infomomenten waar ik mijn onderzoek kon toelichten aan landbouwers en jagers! Ook dank aan het Instituut voor Landbouw-, Visserij- en Voedingsonderzoek (ILVO), Vlaams Instituut voor Technologisch en Onderzoek (VITO) en In2Sky om mij bij te staan in mijn drone-onderzoek.
Special thanks to the jury members Stijn Temmerman, Erik Matthysen, Rudy van Diggelen, Herbert Prins and Oliver Keuling who have provided me with much-appreciated input on my PhD which has certainly improved my thesis.

Mijn vrienden zorgden op de juiste momenten voor ontspanning en afleiding van mijn doctoraat en verdienen hiervoor zeker hun plaats in mijn dankwoord. Allereerst wil ik Michel bedanken. Jij bent mijn beste maatje en staat altijd klaar voor mij, waarvoor eeuwige dank! Van onze epic water fights op kot tot onze vele onvergetelijke reizen naar Frankrijk, Zwitserland, Jordanië, Israël, Hawaië en de vele reizen die nog gaan volgen (als we ooit terug op hetzelfde moment tijd en geld hebben ; ) ), samen zagen over ons doctoraat en ons leven, dromen over onze carrière-ambities, luidkeels meezingen zodat mijn onbetaalde muzikale talenten niet te veel opvallen (remember Chandelelier.. ik denk dat ik erover ben!), mij blijven trainen om chocotoffs te ontwijken, onze vele onbetaalde kajak-tripjes en reisjes naar Parijs, jij maakt mijn leven zoveel leuken, zotter en onbezonnener. Samen met Bernd zijn jullie de perfecte fosterouders, mij voorzien van eten, van een slaapplekje en eigenlijk gewoon van een tweede thuis, jullie zijn de besten!

Verder zijn mijn vriendinnen van de unif Katrien, Anskje, Heleen, Marjolein, Steffi en Rafaëlle één voor één topvrouwen die mijn studententijd een toptijd hebben gemaakt en ik kijk altijd uit naar ons volgende babbelfeestje of etentje. Vivo, you’re the coolest South-African in the world and are such a great friend, together with Leo, Bruno, Janani, Tom and Katherine, you made my masters so special! Although we don’t see each other much, I’m so happy to know you all!

Ook mijn vriendinnen uit Limburg Sanne K, Anneleen C, Ine, Eline, Veronique, Valerie, Lien, Jolien en Sanne Puh zijn maatjes voor het leven. Onze legendarische partypoep-feestjes en -weekendjes, dinner-dates, housewarmings, babyshowers en babyborrels zijn perfecte ontspanningsmomentjes geweest tijdens mijn doctoraat en waren nog vele malen leuken omdat jullie erbij waren! Jullie zijn fantastisch! Inge, Sofie en Dorien, onze ontbijt traditie is iets waar ik altijd naar uit kijk om van iedereen
nieuwtjes te horen en herinneringen op te halen! En Inge, ik kijk er al naar uit
dat we meer gaan kunnen afspreken als we (hopelijk) samen in de VS zijn!

De belangrijksten in mijn leven en degenen die mij het meest gesteund hebben
tijdens mijn doctoraat is ongetwijfeld mijn familie! Mama en papa, jullie zijn
prachtmensen alleen al om mij de kans te geven om mij verder te laten studeren.
Jullie onvoorwaardelijke steun tijdens de zoektocht wat ik met mijn leven wil
doen is van onschatbare waarde. Bedankt om altijd voor mij klaar te staan, voor
de dagelijkse telefoontjes, om naar Antwerpen te komen als ik het even niet zag
zitten, om dag-uitstapjes te doen ter ontspanning en om fantastische reizen naar
Zuid-Afrika en Zanzibar met jullie te kunnen delen! Tom & Sarina, Robin &
Bram, Luna en Justin, jullie zorgen voor zo’n fijne sfeer in ons gezin waardoor
het voor mij altijd fantastisch is om naar huis te komen. Bedankt voor de vele
babbeljess, uitstapjes, shopreisjes, fietstochtjes en steun die ik altijd van jullie
krijg! Liefste Oma, bedankt voor er altijd te zijn voor mij! De kracht die jij
uitstraalt geeft me elke keer weer energie om door te gaan met mijn onderzoek.
Liefste ma, bedankt voor de fijne babbeljess als ik tijdens mijn veldwerk weer
eens even kon binnenspringen! Rowen, jij bent het beste wat mij ooit is
overkomen, meter mogen zijn van zo’n topmeid maakt het leven gewoon
duizend keer beter! één lach van jouw en ik vergeet direct mijn zorgen, één
namiddagje spelen met jouw energieke zelve en ik ben gewoon te moe om nog
maar aan de uitdagingen van mijn doctoraat te kunnen denken :)

Mijn doctoraat is een ongelooftelijke levenservaring geworden en wat het nog
beter maakte was dat elk van jullie er een deel van waren! Bedankt!
References


Alexander, N., Massei, G. and Wint, W. (2016) The European distribution of *Sus scrofa*. Model outputs from the project described within the poster–where are all the boars? An attempt to gain a continental perspective. Open Health Data 4(1).


References


Cahill, S., Llimona, F., Cabañeros, L. and Calomardo, F. (2012) Habituation of wild boar (*Sus scrofa*) to urban areas: traits from Collserola Natural Park (Barcelona) and comparison with other cities. Animal Biodiversity and Conservation 35(2), 221-233.


Frank, B.F. (2012) The conflict-coexistence continuum: a human dimension case study on wild boar management in Italy, Memorial University, St. John’s, Newfoundland, Canada.


References


Keuling, O., Stier, N. and Roth, M. (2008b) Commuting, shifting or remaining? Different spatial utilisation patterns of wild boar Sus scrofa L. in forest and field crops during summer. Mammalian Biology 74, 145-152.


species responses to national forest management. Ecology and evolution 7(1), 368-378.


Salamí San Juan, E., Barrado Muxí, C. and Pastor Llorens, E. (2014) UAV flight experiments applied to the remote sensing of vegetated areas.


Signorille, A.L. and Evans, J. (2007) Damage caused by the American grey squirrel (Sciurus carolinensis) to agricultural crops, poplar plantations and semi-natural woodland in Piedmont, Italy. Forestry 80(1), 89-98.


change risks faced by 90 California vertebrate species of concern. Diversity and Distributions 20(3), 334-343.


Supplementary materials
Chapter 2

Appendix 2.1: Online survey

1. In which agricultural sector is your farm situated?
   a. Cattle farming
   b. Pig farming
   c. Chicken farming
   d. Arable farming
   e. Fruit culture
   f. Vegetable cultivation
   g. Floriculture
   h. Others

2. In which municipality is the largest part of the farm situated?
   Fill in

3. What is the total surface of your farm?
   a. 0 – 10 ha
   b. 10 – 20 ha
   c. 20 – 40 ha
   d. 40 – 60 ha
   e. More than 60 ha

4. Do you have crops which can be sensitive for wild boar damage (grasslands are regarded as crops)?
   a. Yes
   b. No

5. Did you already experience damage by wild boar?
   a. Yes
   b. No

6. Which crops have already been damaged?
   a. Grassland
   b. Wheat
   c. Rye
   d. Barley
   e. Triticale
   f. Oats
   g. Maize
   h. Sugar beets
   i. Potatoes
   j. Fodder beets
   k. Other

7. What is the damaged surface for each of the damaged crops?
   Fill in

8. Did you already take preventive measures to prevent damage from wild boar? And which ones did you take?
   Fill in

9. Did you already request damage compensation?
   a. Yes
   b. No

10. Why did you not request damage compensation?
    Fill in
### 11. Score the following statements

#### a. I would take preventive measures because:

<table>
<thead>
<tr>
<th>I think the damage compensation in Flanders is insufficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>I don’t think the current hunting pressure has a positive effect on the population</td>
</tr>
<tr>
<td>Wild boar numbers and damage will increase</td>
</tr>
<tr>
<td>Investing in preventive measures is advantageous</td>
</tr>
<tr>
<td>I don’t have to worry about future damage</td>
</tr>
<tr>
<td>I can prevent discussions with hunters and other farmers</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I totally agree</th>
<th>I agree</th>
<th>Neutral</th>
<th>I do not agree</th>
<th>I totally do not agree</th>
</tr>
</thead>
</table>

#### b. I don’t take preventive measures because:

<table>
<thead>
<tr>
<th>There are no wild boar in this area</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is no damage in this area</td>
</tr>
<tr>
<td>The investment is too high compared to the damage</td>
</tr>
<tr>
<td>Measures should be taken by hunters, not by farmers</td>
</tr>
<tr>
<td>Measures should be taken by owners of nature areas, not by farmers</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I totally agree</th>
<th>I agree</th>
<th>Neutral</th>
<th>I do not agree</th>
<th>I totally do not agree</th>
</tr>
</thead>
</table>
Appendix 3.1: Overview of fields and photos of damaged fields by wild boar

<table>
<thead>
<tr>
<th>Year</th>
<th>Crop</th>
<th>Number of fields photographed</th>
<th>Total number of photos</th>
<th>Total area photographed (ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>Maize</td>
<td>21</td>
<td>1152</td>
<td>52.54</td>
</tr>
<tr>
<td></td>
<td>Grassland</td>
<td>1</td>
<td>12</td>
<td>0.42</td>
</tr>
<tr>
<td>2016</td>
<td>Maize</td>
<td>26</td>
<td>3212</td>
<td>65.62</td>
</tr>
<tr>
<td></td>
<td>Grassland</td>
<td>33</td>
<td>4856</td>
<td>69.99</td>
</tr>
<tr>
<td>2017</td>
<td>Maize</td>
<td>32</td>
<td>3185</td>
<td>96.54</td>
</tr>
<tr>
<td></td>
<td>Grassland</td>
<td>20</td>
<td>1383</td>
<td>53.09</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>133</td>
<td>14030</td>
<td>338.2</td>
</tr>
</tbody>
</table>
## Appendix 3.2: Mean decreasing accuracies (MDA) of random forest model

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Variable type</th>
<th>MDA in grassland RF-model</th>
<th>MDA in maize field RF-model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel area</td>
<td>Shape</td>
<td>98.94</td>
<td>164.86</td>
</tr>
<tr>
<td>Gray level co-occurrence matrix (GLCM) Angular 2\textsuperscript{nd} moment</td>
<td>Texture</td>
<td>105.52</td>
<td>140.93</td>
</tr>
<tr>
<td>Gray level co-occurrence matrix (GLCM) Contrast</td>
<td>Texture</td>
<td>106.94</td>
<td>133.27</td>
</tr>
<tr>
<td>Gray level co-occurrence matrix (GLCM) Correlation</td>
<td>Texture</td>
<td>104.4</td>
<td>162.97</td>
</tr>
<tr>
<td>Gray level co-occurrence matrix (GLCM) Dissimilarity</td>
<td>Texture</td>
<td>111.84</td>
<td>117.19</td>
</tr>
<tr>
<td>Gray level co-occurrence matrix (GLCM) Entropy</td>
<td>Texture</td>
<td>105.76</td>
<td>139.54</td>
</tr>
<tr>
<td>Gray level co-occurrence matrix (GLCM) Homogeneity</td>
<td>Texture</td>
<td>109.45</td>
<td>134.86</td>
</tr>
<tr>
<td>Gray level co-occurrence matrix (GLCM) Mean</td>
<td>Texture</td>
<td>103.02</td>
<td>113.75</td>
</tr>
<tr>
<td>Gray level co-occurrence matrix (GLCM) Standard Deviation</td>
<td>Texture</td>
<td>110.51</td>
<td>139.85</td>
</tr>
<tr>
<td>Mean difference to Neighbors layer Red - distance 0</td>
<td>Neighbor</td>
<td>120.16</td>
<td>149.27</td>
</tr>
<tr>
<td>Mean difference to Neighbors layer Green - distance 0</td>
<td>Neighbor</td>
<td>173.83</td>
<td>267.11</td>
</tr>
<tr>
<td>Mean layer Red</td>
<td>Spectral</td>
<td>149.46</td>
<td>140.48</td>
</tr>
<tr>
<td>Mean layer Green</td>
<td>Spectral</td>
<td>159.5</td>
<td>140.96</td>
</tr>
<tr>
<td>Mean layer Blue</td>
<td>Spectral</td>
<td>117.54</td>
<td>146.15</td>
</tr>
<tr>
<td>Number of Pixels</td>
<td>Shape</td>
<td>100.43</td>
<td>165.47</td>
</tr>
<tr>
<td>Ratio to scene layer Red</td>
<td>Spectral</td>
<td>137.93</td>
<td>250.67</td>
</tr>
<tr>
<td>Ratio to scene layer Green</td>
<td>Spectral</td>
<td>235.5</td>
<td>183.44</td>
</tr>
<tr>
<td>Ratio to scene layer Blue</td>
<td>Spectral</td>
<td>254.04</td>
<td>240.12</td>
</tr>
<tr>
<td>Relative contribution of layer</td>
<td>Spectral</td>
<td>118.11</td>
<td>166.57</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Red to brightness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative contribution of layer</td>
<td>Spectral</td>
<td>185.7</td>
<td>346.09</td>
</tr>
<tr>
<td>Green to brightness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative contribution of layer</td>
<td>Spectral</td>
<td>185.13</td>
<td>187.9</td>
</tr>
<tr>
<td>Blue to brightness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X distance to left border</td>
<td>Shape</td>
<td>155.09</td>
<td>237.24</td>
</tr>
<tr>
<td>X distance to right border</td>
<td>Shape</td>
<td>111.82</td>
<td>238.29</td>
</tr>
<tr>
<td>Y distance to scene bottom</td>
<td>Shape</td>
<td>118.57</td>
<td>271.68</td>
</tr>
<tr>
<td>border</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y distance to scene left border</td>
<td>Shape</td>
<td>123.39</td>
<td>214.97</td>
</tr>
</tbody>
</table>
Appendix 3.3: Damaged area calculations

Link:  https://wildlife.onlinelibrary.wiley.com/doi/10.1002/wsb.916
Appendix 4.1: Maize model plots

Figure A4.1.1: Relation between holdout deviance and number of trees for A) the landscape model maize model, B) the field specific maize model, C) combined maize model
Figure A4.1.2: Relation between damage probability and all included variables with their importance between brackets for the landscape maize model. Black line represents fitted relationship, red line represents smoothed relationship.

Figure A4.1.3: Relation between damage probability and all included variables with their importance between brackets for the field specific maize model. Black line represents fitted relationship, red line represents smoothed relationship.
Figure A4.1.4: Relation between damage probability and all included variables with their importance between brackets for the combined maize model. Black line represents fitted relationship, red line represents smoothed relationship.
Table A4.1.1: Interactions included in the landscape maize model with interactions sizes and interaction plots.

<table>
<thead>
<tr>
<th>Variable 1</th>
<th>Variable 2</th>
<th>Size</th>
<th>Interaction plot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of grassland</td>
<td>Percentage of maize</td>
<td>7.65</td>
<td><img src="image1" alt="Interaction plot 1" /></td>
</tr>
<tr>
<td>Percentage of grassland</td>
<td>Percentage of low cover</td>
<td>6.76</td>
<td><img src="image2" alt="Interaction plot 2" /></td>
</tr>
</tbody>
</table>
Table A4.1.2: Interactions included in the combined maize model with interactions sizes and interaction plots.

<table>
<thead>
<tr>
<th>Variable 1</th>
<th>Variable 2</th>
<th>Size</th>
<th>Interaction plot</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAO</td>
<td>Percentage of forested edge</td>
<td>5.08</td>
<td><img src="image" alt="3D plot" /></td>
</tr>
<tr>
<td>Percentage of grassland</td>
<td>Percentage of low cover</td>
<td>4.50</td>
<td><img src="image" alt="3D plot" /></td>
</tr>
<tr>
<td>Percentage of low cover</td>
<td>Percentage of forest</td>
<td>4.31</td>
<td><img src="image" alt="3D plot" /></td>
</tr>
<tr>
<td>Percentage of maize</td>
<td>Percentage of urbanization</td>
<td>2.40</td>
<td><img src="image" alt="3D plot" /></td>
</tr>
</tbody>
</table>
Appendix 4.2: Grassland model plots

Figure A4.2.1: Effects plots of final grassland landscape model with increasing damage probability when distance to nearest road (m) increases and distance to nearest forest (m) decreases.

Figure A4.2.2: Effects plots of final field specific grassland model with decreasing damage probability when applying organic fertilization, increasing damage probability when applying inorganic fertilization, decreasing damage probability with increase age when maize was cultivated before and increasing damage probability with increasing age when maize was not cultivated before.
Figure A4.2.3: Effects plots of final field specific grassland model with increasing damage probability when distance to nearest road (m) increases and distance to nearest forest (m) decreases, decreasing damage probability when applying organic fertilization, increasing damage probability when applying inorganic fertilization, decreasing damage probability with increase age when maize was cultivated before and increasing damage probability with increasing age when maize was not cultivated before.
Appendix 4.3: MESS maps

Figure A4.3.1: MESS maps (multivariate environmental similarity surface) for landscape model of A) maize fields and B) grasslands with positive MESS-values (green to red) reflecting full range of variable conditions are included in the data (higher positive values are better represented by the data) and negative MESS-values (blue) reflecting variable conditions which are not included in the data thus the model is extrapolating.
# Chapter 5

## Appendix 5.1: Reclassification of CORINE land cover classes

<table>
<thead>
<tr>
<th>CLC_CODE</th>
<th>Class CORINE</th>
<th>Label CORINE</th>
<th>Reclassified land cover type</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>Artificial surfaces</td>
<td>Continuous urban fabric</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>112</td>
<td>Artificial surfaces</td>
<td>Discontinuous urban fabric</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>121</td>
<td>Artificial surfaces</td>
<td>Industrial or commercial units</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>122</td>
<td>Artificial surfaces</td>
<td>Road and rail networks and associated land</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>123</td>
<td>Artificial surfaces</td>
<td>Port areas</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>124</td>
<td>Artificial surfaces</td>
<td>Airports</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>131</td>
<td>Artificial surfaces</td>
<td>Mineral extraction sites</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>132</td>
<td>Artificial surfaces</td>
<td>Dump sites</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>133</td>
<td>Artificial surfaces</td>
<td>Construction sites</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>141</td>
<td>Artificial surfaces</td>
<td>Green urban areas</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>142</td>
<td>Artificial surfaces</td>
<td>Sport and leisure facilities</td>
<td>Urban Coverage</td>
</tr>
<tr>
<td>211</td>
<td>Agricultural areas</td>
<td>Non-irrigated arable land</td>
<td>Agricultural Coverage</td>
</tr>
<tr>
<td>222</td>
<td>Agricultural areas</td>
<td>Fruit trees and berry plantations</td>
<td>Agricultural Coverage</td>
</tr>
<tr>
<td>231</td>
<td>Agricultural areas</td>
<td>Pastures</td>
<td>Agricultural Coverage</td>
</tr>
<tr>
<td>242</td>
<td>Agricultural areas</td>
<td>Complex cultivation patterns</td>
<td>Agricultural Coverage</td>
</tr>
<tr>
<td>243</td>
<td>Agricultural areas</td>
<td>Land principally occupied by agriculture, with significant areas of natural vegetation</td>
<td>Agricultural Coverage</td>
</tr>
<tr>
<td>311</td>
<td>Forest and semi natural areas</td>
<td>Broad-leaved forest</td>
<td>Forest Coverage</td>
</tr>
<tr>
<td>312</td>
<td>Forest and semi natural areas</td>
<td>Coniferous forest</td>
<td>Forest Coverage</td>
</tr>
<tr>
<td>313</td>
<td>Forest and semi natural areas</td>
<td>Mixed forest</td>
<td>Forest Coverage</td>
</tr>
<tr>
<td>321</td>
<td>Forest and semi natural areas</td>
<td>Natural grasslands</td>
<td>Low Coverage</td>
</tr>
<tr>
<td>322</td>
<td>Forest and semi natural areas</td>
<td>Moors and heathland</td>
<td>Low Coverage</td>
</tr>
<tr>
<td>324</td>
<td>Forest and semi natural areas</td>
<td>Transitional woodland-shrub</td>
<td>Low Coverage</td>
</tr>
<tr>
<td>331</td>
<td>Forest and semi natural areas</td>
<td>Beaches, dunes, sands</td>
<td>Low Coverage</td>
</tr>
<tr>
<td>411</td>
<td>Wetlands</td>
<td>Inland marshes</td>
<td>Low Coverage</td>
</tr>
<tr>
<td>412</td>
<td>Wetlands</td>
<td>Peat bogs</td>
<td>Low Coverage</td>
</tr>
</tbody>
</table>
Appendix 5.2: DAPC output

Figure A5.2.1: Relation between the number of clusters and the BIC-value. Based on this graph, 4 clusters are chosen for further analysis.

Explanations on the choice of the number of clusters based on the Jombart and Collins (2015):

As there is no ‘true number of clusters’, the best number of clusters can be normally be distinguished by a clear transition of BIC decrease towards BIC increase. In practice, this change is often unclear showing that reality is more complex than some clear-cut isolated populations. This is indeed the case in this analysis towards wild boar clusters in Flanders. We see a steep decrease of BIC until a cluster number of 4 after which there is a transition period until a cluster number of ±25. In these cases, ‘the actual number retained is merely a question of personal taste’ (Jombart and Collins, 2015). We have therefore chosen 4 clusters to be able to make the comparison with STRUCTURE-analyses possible.
Figure A5.2.2: The Comparison of STRUCTURE (same figure as figure 2F in the chapter 5 without Dutch and German samples and without varying pie chart size) and DAPC clustering results of all 838 Flemish samples with red representing the eastern population (EP), blue representing the western population (WP), yellow representing north-western population (NWP) and green representing the additional DAPC cluster.
Appendix 5.3: SNP dataset

Link: http://www.wildlifebiology.org/appendix/wlb-00542
Appendix 5.4: Structure admixture results

Figure A5.4.1: Admixture per section with Western population (WP, blue), Eastern population (EP, red) and North-Western population (NWP, yellow). Sections A) 2007-2009, B) 2009-2011+Dutch and German samples, C) 2011-2013, D) 2013-2015, E) 2015-2016 are shown. Each individual is represented by a separate bar with the q-value describing the proportion of each individual’s ancestry to one of the three populations.
Appendix 5.5: Observed versus predicted genetic distances

Figure A5.5.1: The relationship of observed versus predicted genetic distances (Rousset’s $a$ value) in relation to forest coverage (%).
Appendix 6.1: Multicollinearity

Figure A6.1.1: Correlation matrix of all variables showing R-values of the Pearson Correlation coefficient.
Appendix 6.2: Selection sampling effort

**Concept**

Citizen scientists who record occurrences of certain species in an online portal like waarnemingen.be are motivated by an interest in certain species or species groups. Where they look for species depends on accessibility of certain locations, known locations where a species occurs/can occur etc. These aspects result in observer specific characteristics and in characteristics of their observations allowing to assess a general type of sampling effort by certain groups of observers (i.e. mammal specialist, butterfly specialist, generalist observer, etc.). By defining a typical wild boar observer and their observations we can assess sampling effort for wild boar in Flanders. We assume that wild boar observers register wild boar observations and tracks whenever they encounter them. Where they have been in the landscape can be reconstructed from all observations (of any species) they have reported, thus giving a good representation of sampling effort and thereby allowing to correct for sampling bias in a MaxEnt model.

**Step 1: define wild boar observer**

Based on expert knowledge on observer characteristics of observers reporting records to waarnemingen.be, we decided to choose two different selections to define wild boar observers focusing of observers who enter mammal recordings (excluding bats as these observers are often considered as a separate observer type, because of the specific effort needed to record bats).

2 definitions:

- **Wild boar expert**: an observer has recorded a wild boar on at least 5 separate days
- **Mammal expert**: an observer has recorded a wild boar on 2 at least separate days + observer has recorded 5 other mammal species (not including bat species)
Step 2: define observations which are included in sampling effort

By defining a wild boar observer according to one of the criteria explained above, a certain number of observers are selected. From these observers, we next selected which kind of their observations are included to define sampling effort for wild boar. We made two selections, including different species records. For each, we can count the total number of days for which observations are submitted in a UTM 1x1 km grid cell (called grid cell-day-visit).

2 criteria defining a grid cell-day-visit:

- A selected wild boar observer who recorded a mammal (not including bat species) in a UTM-grid cell
- A selected wild boar observer who recorded a species (all possible plant or animal species) in a UTM-grid cell

Step 3: construct sampling bias file

All possible combinations of our selections result in 4 different raster files describing the number of ‘daghoktellingen’ per UTM-grid. UTM-grids that were not visited (and thus have a sampling effort of 0) were given a value of 0.1 defining non-sampling effort for a first set of the 4 different files and a value of 0.01 for a second set of these different files according to Kramer-Schadt et al. (2013) to assess the sensitivity of sampling bias files. This resulted in a total of 8 sampling bias files (table A6.2.1).
Table A6.2.1: All selections describing sampling effort selections.

<table>
<thead>
<tr>
<th>Selection</th>
<th>Selected observer</th>
<th>Selected DHB</th>
<th>Non-sampling effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection 1.1</td>
<td>Wild boar expert</td>
<td>Mammal species</td>
<td>0.1</td>
</tr>
<tr>
<td>Selection 2.1</td>
<td>Wild boar expert</td>
<td>Mammal species</td>
<td>0.01</td>
</tr>
<tr>
<td>Selection 1.2</td>
<td>Wild boar expert</td>
<td>All species</td>
<td>0.1</td>
</tr>
<tr>
<td>Selection 2.2</td>
<td>Wild boar expert</td>
<td>All species</td>
<td>0.01</td>
</tr>
<tr>
<td>Selection 1.3</td>
<td>Mammal expert</td>
<td>Mammal species</td>
<td>0.1</td>
</tr>
<tr>
<td>Selection 2.3</td>
<td>Mammal expert</td>
<td>Mammal species</td>
<td>0.01</td>
</tr>
<tr>
<td>Selection 1.4</td>
<td>Mammal expert</td>
<td>All species</td>
<td>0.1</td>
</tr>
<tr>
<td>Selection 2.4</td>
<td>Mammal expert</td>
<td>All species</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Step 4: comparing MaxEnt models

For each selection, a MaxEnt model was run using background data corrected according to the sampling bias file connected to a certain selection (areas with higher sampling effort have a higher chance to be selected as background data). Optimal settings were subsequently defined as explained in the methods section. From each MaxEnt model, projections of habitat suitability were made and the AUC-values of model performance based on model validation using hunting bag records were calculated (Table A6.2.2).

We found minimal variation in model performances between different selections and different values for non-sampling effort. These minimal variations were reflected in very similar projections of habitat suitability in Flanders. The sampling bias file of the best performing model (selection 2.1) including Wild boar experts, mammal species observation defining ‘grid cell-day-visit’ and a non-sampling effort of 0.01 was further used.
Table A6.2.2: Background selection, AUC-value and model projections for each of eight selections.

<table>
<thead>
<tr>
<th>Model selection</th>
<th>Background data</th>
<th>AUC -value</th>
<th>Projections to Flanders</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>![Image 1]</td>
<td>0.782</td>
<td>![Image 2]</td>
</tr>
<tr>
<td>2.1</td>
<td>![Image 3]</td>
<td>0.783</td>
<td>![Image 4]</td>
</tr>
<tr>
<td>1.2</td>
<td>![Image 5]</td>
<td>0.766</td>
<td>![Image 6]</td>
</tr>
<tr>
<td>2.2</td>
<td>![Image 7]</td>
<td>0.762</td>
<td>![Image 8]</td>
</tr>
<tr>
<td>1.3</td>
<td>![Image 9]</td>
<td>0.774</td>
<td>![Image 10]</td>
</tr>
<tr>
<td>2.3</td>
<td>![Image 11]</td>
<td>0.776</td>
<td>![Image 12]</td>
</tr>
<tr>
<td>1.4</td>
<td>![Image 13]</td>
<td>0.774</td>
<td>![Image 14]</td>
</tr>
<tr>
<td>2.4</td>
<td>![Image 15]</td>
<td>0.774</td>
<td>![Image 16]</td>
</tr>
</tbody>
</table>
Appendix 6.3: Model response curves
Figure A6.3.1: Response curves of environmental variables for the eight MaxEnt models.
Appendix 6.4: Model projections

Figure A6.4.1: All model projection of habitat suitability to Flanders. Red colours indicate high habitat suitability; green colours indicate low habitat suitability.
Appendix 6.5: Maximum-Minimum habitat suitability

Figure A6.5.1: Maximum and minimum habitat suitability from all 8 MaxEnt models combined.

Table A6.5.1: Suitable habitat (habitat suitability < 0.50) and unsuitable habitat (habitat suitability < 0.5) in Flanders according to minimum and maximum habitat suitability from all 8 MaxEnt models combined.

<table>
<thead>
<tr>
<th></th>
<th>Suitable habitat</th>
<th>Unsuitable habitat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum suitability</td>
<td>19.21 %</td>
<td>80.79 %</td>
</tr>
<tr>
<td>Minimum suitability</td>
<td>7.67 %</td>
<td>92.33 %</td>
</tr>
</tbody>
</table>
Appendix 6.6: MESS maps

Figure A6.6.1: MESS maps (multivariate environmental similarity surface) with positive MESS-values (green to red) reflecting full range of variable conditions are included in the data and negative MESS-values (blue) reflecting variable conditions that are not included in the data thus the model is extrapolating.