Analyses of road-kills data harvested voluntarily: the case of the Eurasian badger (Meles meles), the European polecat (Mustela putorius) and the raccoon (Procyon lotor) in Wallonia.

Auteur : Ziegler de Ziegleck aùf Rheingrüb, Henry
Promoteur(s) : Licoppe, Alain
Faculté : Gembloux Agro-Bio Tech (GxABT)
Diplôme : Master en bioingénieur : gestion des forêts et des espaces naturels, à finalité spécialisée
Année académique : 2016-2017
URI/URL : http://hdl.handle.net/2268.2/3110

ANALYSES OF ROAD-KILLS DATA HARVESTED VOLUNTARILY: THE CASE OF THE EURASIAN BADGER (*MELES MELES*), THE EUROPEAN POLECAT (*MUSTELA PUTORIUS*) AND THE RACCOON (*PROCYON LOTOR*) IN WALLONIA.

HENRY ZIEGLER DE ZIEGLECK AÙF RHEINGRÜB

TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE MASTER BIOINGÉNIEUR EN GESTION DES FORÊTS ET DES ESPACES NATURELS.

ANNÉE ACADÉMIQUE 2016-2017

PROMOTEUR: ALAIN LICOPPE

ANALYSES OF ROAD-KILLS DATA HARVESTED VOLUNTARILY: THE CASE OF THE EURASIAN BADGER (*MELES MELES*), THE EUROPEAN POLECAT (*MUSTELA PUTORIUS*) AND THE RACCOON (*PROCYON LOTOR*) IN WALLONIA.

HENRY ZIEGLER DE ZIEGLECK AÙF RHEINGRÜB

TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE MASTER BIOINGÉNIEUR EN GESTION DES FORÊTS ET DES ESPACES NATURELS

ANNÉE ACADÉMIQUE 2016-2017

PROMOTEUR: ALAIN LICOPPE

ACKNOWLEDGMENT

This master thesis has been conducted in partnership with the Zoogeography Research Unit of the University of Liège and the Département d'Étude du Milieu Naturel et Agricole (DEMNA).

Many thanks to Alain Licoppe and Vinciane Schockert for having supervising this master thesis.

Many thanks as well to my family and my friends for their support, opinions and advice.

Finally, special thanks to my father for having writing the program to sort double-counts out.

Please note that English is not my mother tongue. Despite lots of efforts, some language mistakes could remain in the text.

To ask questions or make comments about this master thesis, you are very welcome to contact me at
henry.ziegler@posteo.net

ABSTRACT

Road infrastructures impact wildlife species in several ways. In particular, animals could suffer from habitat loss, traffic mortality, barrier to movements and populations subdivision. Despite the importance of those issues and the large amount of available data, only few published papers investigated road-wildlife interactions in Wallonia, Belgium so far. Road-kills observations recorded between 2006 and 2016 in three different databases have therefore been analysed, focusing on the Eurasian badger (*Meles meles*), the European polecat (*Mustela putorius*) and the raccoon (*Procyon lotor*). Those data have been harvested voluntarily, without any sampling program. In this master thesis, an original methodology has first been designed to sort out and select relevant data. To do so, distances to roads have been used and double-counts have been removed. Then, the goal was to test whether roadkill observations could serve as an indicator of wildlife populations' status and trends. Results showed poor to medium similarity between collision data and data that were considered as the reference. Finally, impacts of roads on these mammals have been investigated. Hotspots maps have been drawn, and it has been calculated that at least 7,42% to 13,69% of badgers' populations are killed on roads each year. However, all those results are only partially reliable as they depend on search effort, which was unknown. For further research, it is recommended to improve data encoding and to investigate deeper the issue of sampling effort.

RÉSUMÉ

Les infrastructures routières impactent la faune sauvage de différentes manières, notamment par la perte d'habitat, la mortalité routière, l'effet de barrière et la subdivision des populations. Malgré l'importance de cette problématique et la grande quantité de données disponibles, seules quelques études se sont intéressées jusqu'ici aux interactions entre la faune et les routes en Wallonie (Belgique). Les données de mortalité routière encodées entre 2006 et 2016 dans trois différentes bases de données ont donc été analysées, en se concentrant sur le blaireau (*Meles meles*), le putois (*Mustela putorius*) et le raton-laveur (*Procyon lotor*). Ces données ont été récoltées sur une base volontaire, sans aucun programme d'échantillonnage. Dans ce travail de fin d'études, une méthode originale a d'abord été élaborée pour trier les données et sélectionner celles qui étaient pertinentes. Pour ce faire, les distances aux routes ont été utilisées et les double-comptages ont été retirés. Ensuite, le but était de voir si les observations d'animaux tués sur les routes pouvaient servir d'indicateur des populations. Des concordances faibles à moyennes ont été trouvées entre les données de mortalité et les données servant de référence pour les populations. Finalement, l'impact des routes sur ces mammifères a été étudié. Des cartes de « hotspots » ont été établies et il a été calculé qu'au minimum entre 7,42% et 13,69% de la population de blaireau était tuée chaque année sur les routes. Cependant, tous ces résultats sont seulement en partie fiables étant donné qu'ils dépendent de l'effort d'échantillonnage, qui était inconnu. Pour de prochains travaux, il est suggéré d'améliorer l'encodage des données et d'étudier plus en détail le problème de l'effort d'échantillonnage.

TABLE OF CONTENT

1 / INTRODUCTION

1.1 / Roads and wildlife

Roads impact ecosystems in several ways. They can lead to changes in hydrology and water quality, release of chemical pollutants, increase in noise and light, but roads are also a major source of mortality and fragmentation for many animal species (Coffin, 2007). For example, in the United States, it seems that traffic kills more vertebrates every year than hunting (Forman et al., 1998). This issue has raised concerns among scientists and activists for past decades, leading to a large amount of studies and actions in many countries.

Jaeger et al. (2005) reviewed ways in which roads affect wildlife populations, and identified four main effects:

- Habitat loss. Habitat is directly changed when building roads and their verges. Habitat quality can also be reduced in the surroundings of roads because of noise, light and pollutants emissions.
- Traffic mortality. Animals trying to cross roads are exposed to risks of collision with vehicles.
- Barrier to movements. Because animals are afraid to cross tarmac or because they are killed when doing so, roads can *de facto* act as barriers for wildlife. It therefore reduces access to resources, food, mates or habitat patches. For example, Rondinini et al. (2002) found out that large roads are considered as barriers by hedgehogs.
- Population subdivision. Because animals are restricted in their movements, less migration can occur. Roads can therefore separate and isolate subpopulations and thus affect genetic exchanges. For example, Frantz et al. (2012) showed that a highway could be a gene flow barrier for red deer.

All those aspects are likely to reduce populations' survival rates and persistence (Jaeger et al., 2005).

Factors influencing occurrence of wildlife-vehicles collisions are numerous and species-specific (Gunson et al. 2011). According to the summary established by Grilo et al. (2009), "temporal variations in road-kill are related to differences in species behaviour and activity, e.g., foraging and resting, mating, and dispersal of juveniles" whereas spatial variations "appear to depend on population density, species biology, habitat and landscape structure, as well as road and traffic characteristics" (Grilo et al., 2009). Every situation is therefore different and requires specific approaches (Morelle et al., 2013).

As wildlife-vehicles collisions have strong impacts on wildlife but also on drivers (it can cause injuries, death and damages to vehicles and infrastructures), mitigation measures have been designed for decades to try to reduce those problems.

Fencing is a common measure undertaken to decrease the number of collisions with large mammals. Thus, it also increases drivers' safety: by fitting out fences along roads, animals are not able to cross them anymore, which therefore significantly reduces collisions risks. However, those infrastructures often enhance roads' barrier effects (Jaeger et al., 2004). As a consequence, crossing structures are needed to ensure transportation network permeability to wildlife. Many types of passes exist, namely pipe culverts, box culverts, tunnels, underpasses (associated or not to water courses), overpasses, and so on (Glista et al., 2009). The latter are bridges covered with soil and vegetation in order to restore ecological connectivity. They are generally designed for large species (Glista et al., 2009) (Figure 1).

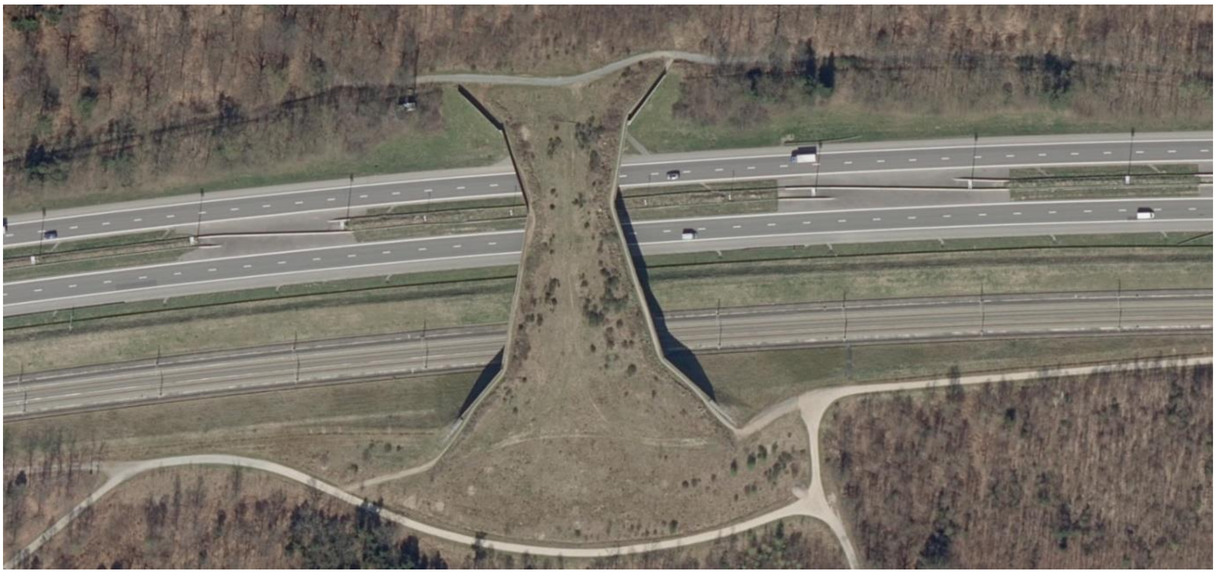


Figure 1 : Overpass for wildlife at Grunhaut, Wallonia, Belgium. Data source : Service Public de Wallonie (SPW), 2015.

However, fencing is usually only possible along major roads. Then other techniques also exist, such as the use of reflectors, wildlife-crossing signs or traffic speed reductions (van Langevelde et al., 2009). Nevertheless, despite the growing amount of studies focusing on all those mitigation measures, their effectiveness sometimes remains unclear or even unknown (Glista et al., 2009; Rytwinski et al., 2016).

Apart from studying impacts of roads on wildlife populations, some researchers also suggested that roadkill data could be used to monitor wild animals' populations, in particular mammals (For example Gehrt, 2002 ; Baker et al., 2004; George et al., 2011). Indeed, wild mammals are usually elusive, discreet and nocturnal, so that observing them directly is very difficult. Monitoring programmes therefore have to rely on indirect indices such as tracks, faeces or game bags (George et al., 2011). On the opposite, carcasses along roads are easy to observe. What if they could indicate populations' abundance and trends?

But to serve as an indicator, a clear relationship between roadkill data and actual populations' status has to exist (Baker et al., 2004). Studies' conclusions are various. Baker et al. (2004) and George et al. (2011) found a direct relationship, while Rolley et al. (1992), Gehrt (2002) and Seiler (2004) found a variable reliability of this kind of indicator. However, they all concluded that such an index would only work using a large spatial and temporal scale.

1.2 / Existing literature about road-wildlife interactions in Wallonia

Although road-wildlife interaction is a very well documented topic that has been investigated for decades all around the world, only few studies on this theme have been conducted so far in Wallonia (southern administrative region of Belgium). Available literature has been here reviewed (Table 1):

Table 1 : Existing body of literature about road-wildlife interactions in Wallonia, Belgium. This list is as much comprehensive as possible. However, it is not guaranteed that it is exhaustive.

Reference	Type of document	Content summary
Holsbeek et al. (1999)	Scientific paper	Data harvested by volunteers, mainly in Flanders, Belgium. Hedgehogs and rabbits accounted for the majority of the carcasses. A rough estimate of 4 000 000 animals being killed each year in Belgium has been calculated
Frantz et al. (2012)	Scientific paper	Investigates genetic structures induced in red deer and wild boar populations isolated by a highway in southern Beligum.
Morelle et al. (2013)	Scientific paper	Paper focused on ungulates. Analyses spatial and temporal patterns of collisions thanks to data recorded by the police when accidents occur(covers the period 2003 - 2011 and accounts for 3965 accidents with wildlife)
Lehaire et al. (2013)	Popular scientific article	Popular scientific article that summarizes the paper of Morelle et al. for the public.
Baghli (1996)[1]	Master thesis	Master thesis focused on overpasses, especially the one located at Rulles (Ardennes, southern Wallonia).
Maron (2013)[2]	Master thesis	Master thesis associated to the paper of Morelle et al. Investigate roadkill aggregations, temporal patterns and factors influencing collisions with ungulates.
Denys (2017)[3]	Master thesis	Summarizes impact of transportation infrastructures on wildlife. Then tries to determine optimal location for fauna passages thanks to GIS modelling.

[1] Baghli M., 1996. *Etude de passages pour la grand faune: cas du passage à gibier de la forêt de Rulles (Ardennes)*. Travail de fin d'études: FUCAM (Belgique).
[2] Maron J., 2013. *Analyse spatio-temporelle des collisions routières impliquant des ongulés sauvages en wallonie*. Travail de fin d'étude: Gembloux Agro-Bio Tech, Université de Liège (Belgique).
[3] Denys M., 2017. *Atténuation des impacts des infrastructures de transport sur la faune terrestre : déterminer la localisation optimale des passages fauniques grâce aux Systèmes d'information géographique (SIG)*. Travail de fin d'études: Institut de Gestion de l'Environnement et d'Aménagement du Territoire. Université Libre de Bruxelles (Belgique).

As a result, only few facts and figures are available in Wallonia. Considering this little number of studies on this topic, the importance of road impacts on populations and the potential great amount of unexploited data, it has been decided to focus this master thesis on wild mammals mortality related to roads in Wallonia

More details will now be given about Walloon road network, available data and databases, and species that will be at the centre of this thesis.

1.3 / Road network in Wallonia

Compared to other European regions, Wallonia has a high population density (212 inhabitants/km²) which also leads to one of the densest road network of Europe (on average 4,82 km of road per km² in Wallonia in 2010; 1,10 km/km² in the European Union)(SPW, 2014). There are three road categories: highways (869 km), regional roads (7583 km) and municipal roads (49 189 km) (Figure 2). Those figures were calculated in 2010, but road infrastructures are well developed and haven't changed for past years. On the opposite, the amount of traffic has significantly increased and is still increasing (+95% on highways and +38% on regional roads between 1990 and 2010)(SPW, 2014).

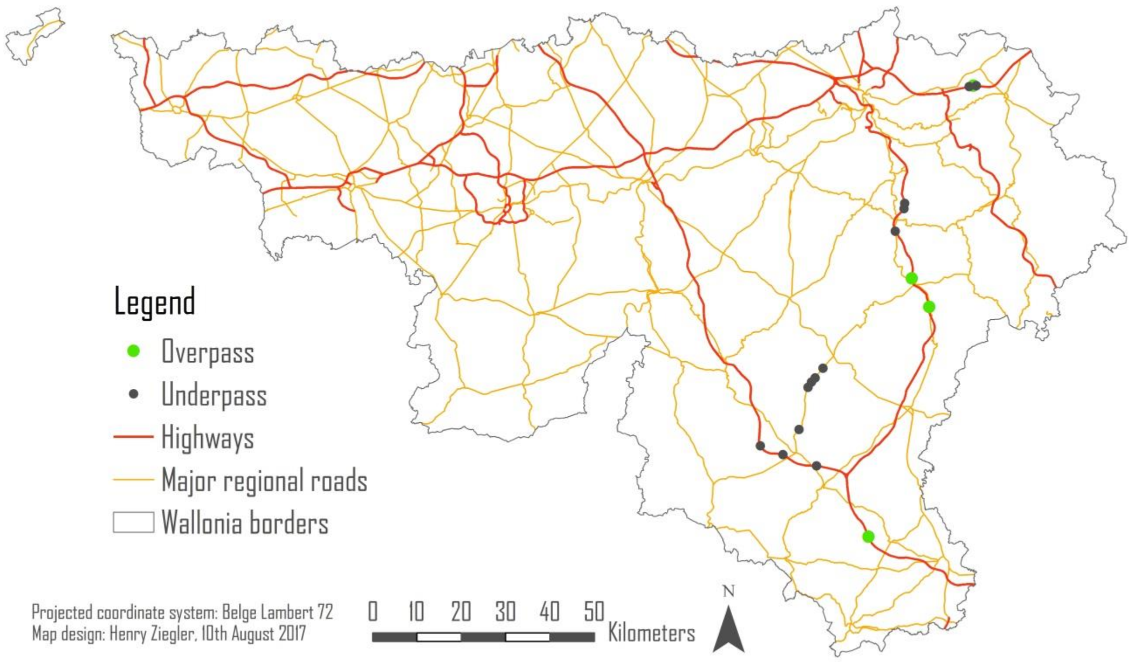


Figure 2 : Highways and major regional road network of Wallonia. Under and overpasses are located as well.

Data about mitigation measures fitted out in Wallonia are not easily accessible, and scientists often miss accurate information about it. I therefore contacted several employees at road administration to gather together available information in order to produce an original summary. The main results are presented here:

- Only highways are fenced, and exclusively along sections that go through forested areas. Those fences are designed to prevent ungulates from crossing the road and causing collisions. It seems that no map of their exact location exists.

- However, those fences have often been fitted out 40 years ago. It is likely that many of them suffer from damages and lack of maintenance. Their efficiency is not ensured.

- 13 underpasses and 4 overpasses are recorded (see Figure 2). They are listed in the Appendix 1. Aerial views of the four overpasses are also presented there. This list has been established by merging information available at road administration and in the master thesis of Baghli M. (1996) cited above. Please note that administration's database is not always comprehensive and that other unrecorded passes could exist.

- Effectiveness of those passes is uncertain. Up-to-date data about their use are lacking. They have often been build 40 years ago, when knowledge about passes engineering was poor. It is likely that 3 of the four overpasses are too narrow to suit to large ungulates' behaviour. Moreover, many passes are mixed passes (can be used by wildlife and by vehicles as well). Their efficiency for wildlife crossings is questionable.

1.4 / Data and data sources

To achieve this master thesis, I didn't conduct any field work on my own to harvest roadkill data. Instead, I used existing datasets. Indeed, a lot of roadkill observations have been gathered in Wallonia for the last decade and only few studies attempted to analyse them so far (see literature review above). This master thesis was therefore a good opportunity to make this large amount of data profitable.

Unfortunately, there is no centralized database yet to record road-kills observations in Wallonia. Data are scattered between many different actors such as naturalist associations, scientific staffs in universities, nature departments of Wallonian administration, the police, insurance companies, hunters, and so on (Morelle et al., 2013).

Three sources of data have been used here:

- Database of the website www.observations.be
- Walloon administration database (OFFH)
- Dissection data from the Zoogeography Research Unit of the University of Liège.

These databases probably represented the largest amount of available information. Moreover, they have been running for at least ten years and it is very likely that they will continue to be supplied with observations for the coming years. It was therefore relevant to use them. Here are more details about each of these databases:

1.4.1 / Observations.be

Observations.be[4] is a website designed in 2006 by naturalist associations Aves-Natagora and Natuurpunt. It is the Belgian version of www.observation.org. The aim of this website is to encode, manage and share data about biodiversity in Belgium. For this purpose, observations.be provides tools and interfaces to encode observations into a database[5]. The website uses participative-science principles: everyone (professional or not) has the opportunity to use this platform for free and to supply it with data and observations about all kinds of species (plants, insects, birds, mammals, etc.) he/she has seen during his/her daily activities. It means that data are harvested on a voluntary basis, that they are sporadic and are not the result of a coordinated sampling scheme. Whenever possible, data are validated by a group of expert in order to improve database quality. After 10 years, more than 24 million of various observations have been encoded.

When encoding, users are required to provide some details like species, date and number of individuals. They also have to localise their observations on a map thanks to a Google Maps interface. Lambert Belge 72 and WGS84 coordinates are then automatically generated. Moreover, users are given the possibility to add optional details. For example, they can specify animal's behaviour, biotope where it has been seen, the precision degree of the location (10, 100 or 1000 meters) and write some comments about the observation. A screenshot of the form is shown below (Figure 3).

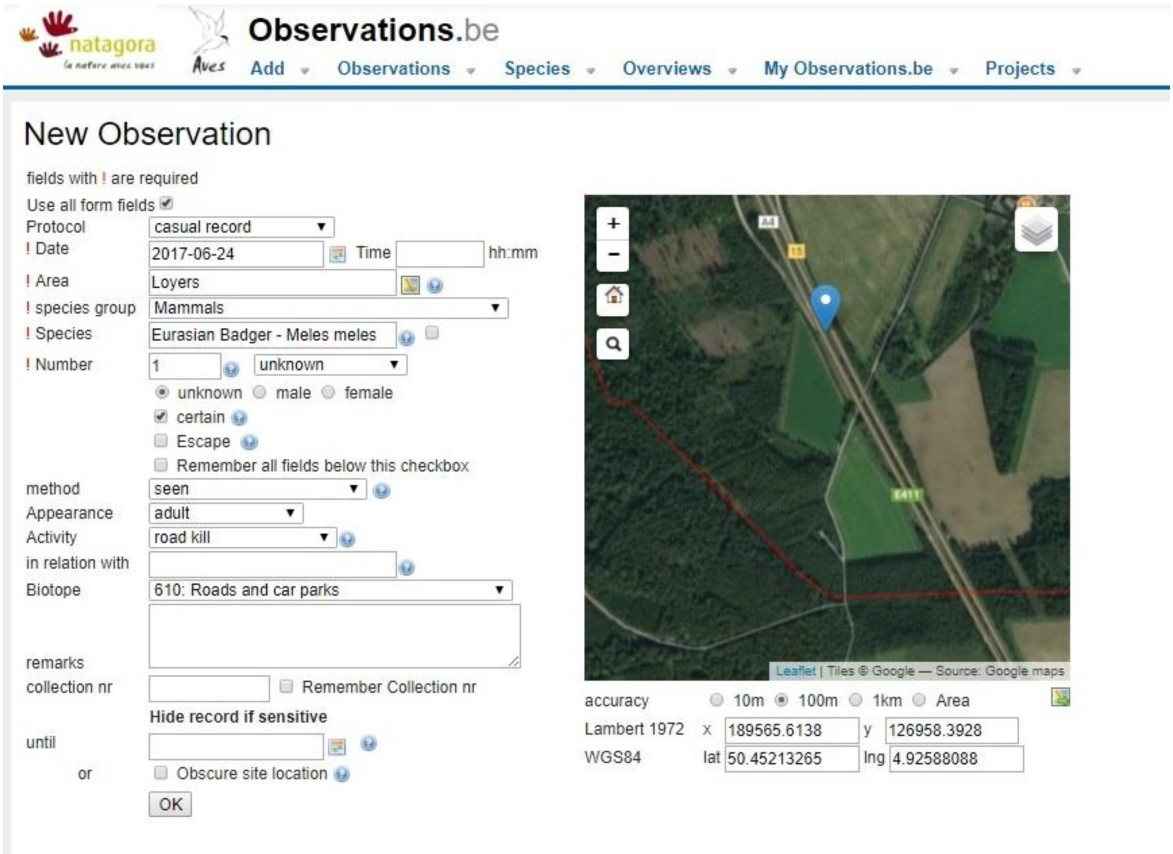


Figure 3: Screenschot of the form of observations.be

---

[4] Link : https://observations.be/

[5] Source : https://observations.be/info.php

1.4.2 / OFFH

L'Observatoire de la Faune, de la Flore et des Habitats (OFFH) is a regional platform that has been set up by the Walloon administration in order to coordinate biological data harvesting, analysis and sharing[6]. It is managed by the department for agricultural and natural areas studies (DEMNA – Département d'Étude du Milieu Naturel et Agricole).

OFFH database contains lots of various flora and fauna observations. Contrary to observations.be, these data are only gathered by and accessible to employees of the Walloon administrations, like foresters or scientific staff members. Observations are harvested during their daily field activities on a voluntary basis. They are mainly opportunistic: staff members have the possibility to encode their observations in OFFH database, but they are not forced to do it. Here again, there is therefore no coordinated sampling programme.

OFFH also contains observations gathered in 2015 and 2016 for a project called Go For Blaireau ("Go For Badger"). The aim was to involve some scientists and naturalists in a regular survey of dead animals (especially badgers but other species were also recorded) along roads. Those people choose a transect they drove regularly (route from home to work for example) and agreed to pay attention to carcasses and to record them. Nevertheless the success of this project has been limited so far[7].

OFFH database works quite as observations.be: employees can provide lots of details about their observations and geolocalise them thanks to an integrated Google Maps interface. Compared to observations.be, OFFH database has the advantage to be fed by professionals. Thus, data quality (in particular species determination or localisation) is expected to be better.

1.4.3 / Dissection data

Other data were also used, in particular those from the Zoogeography Research Unit of the University of Liège, that is in charge of the monitoring of several mammals in Wallonia. Whenever possible, animal carcasses found by foresters or administration's scientific staffs are sent to the Zoogeography lab to carry out autopsies. Many of those animals are roadkill. Normally, dates and precise locations (with associated precisions) where carcasses have been found are noted, but this information is sometimes missing. Again, all these data are sporadic as there is no plan to find and collect carcasses.

[6] Source : http://biodiversite.wallonie.be/fr/l-observatoire-de-la-faune-de-la-flore-et-des-habitats-en-wallonie-ministere-de-la-region-wallonne-dgrne.html?IDC=122

[7] Source : Schockert V., Lambinet C., Libois R., 2016. *Convention SPW/ULg sur 15 espèces de mammifères protégés ou concernés par la Convention de Berne et sur 2 espèces envahissantes. Rapport d'activités 2016.* Liège : Unité de recherche zoogéographique, Université de Liège. (internal report, unpublished)

Consequently, almost all of those data have been harvested on a whole region, without sampling programme, and many of them were gathered by non-professionals. This is not a usual situation: most of road ecology studies are conducted with rigorous protocols by trained scientists (for example: Clevenger et al., 2003; Grilo et al., 2009; Haigh, 2012; and so on).

However, those professional studies are very expensive and time-consuming. They are therefore often restricted to small areas or to some road segments, and are not conducted on the long term.

On the other hand, it is nowadays more and more common to launch citizen science projects to “go beyond data scarcity” (Devictor et al., 2010). Those projects allow conducting long-term studies at low costs. In the case of road ecology, road-kills can therefore be monitored for a long time at a large scale (Vercayie, 2015). This is one of the aims of observations.be. In Flanders, roadkill data recorded in this database have already been used for several years, in particular thanks to the project “Dieren Onder de Wielen” (Vercayie, 2015).

However, many potential sources of errors or biases are associated to data harvested on a voluntary basis, especially when they are gathered by non-professionals. This point will be discussed with more details latter in this report (see paragraph 3.1). Nevertheless some of those sources of errors can already be cited:

- Species misrecognition
- Observations misplaced or with a poor precision
- Wrong description of the state of the animal (road-kills are not encoded in the correct category)
- Duplicates: a same animal could have been encoded in several databases.
- Double-counts: as carcasses are not removed from the road after being detected, a same carcass could have been observed and encoded by different people.

Using those observations is therefore challenging and an appropriate methodology has to be applied to deal with them.

1.5 / Study species

Most of the studies conducted in Wallonia have been focused on ungulates (mainly red deer (*Cervus elaphus*), roe deer (*Capreolus capreolus*) and wild boar (*Sus scrofa*)). However, other species such as badger are also known to be frequent traffic victims (Dekker et al., 2010). Thus, it would be interesting to obtain facts and figures about them. This master thesis was therefore focused on three species, namely the Eurasian badger, the European polecat and the raccoon. More details about these species are given here bellow.

1.5.1 / Eurasian Badger[8]

Eurasian badger (*Meles meles*) is a mustelid characterized by its large size (adults can be 70cm to 90cm long and weigh 9kg to 17kg) and its white head with two sharp black bands (Figure 4).

Figure 4 : A Eurasian badger. "badger" by Hehaden. Published on Flickr under CC BY-NC 2.0 licence.

Badgers mostly live in deciduous forests and forests edges. They can also settle in open environments such as hedged farmlands and heathlands, provided soils are soft enough to dig burrows. They live in social clans.

Badger was largely distributed in Wallonia in the past, but populations seriously dropped down in the eighties because of rabies propagation and burrows gassing. Since foxes have been vaccinated and burrows gassing has been stopped, number of individuals has increased. Badgers are nowadays widely spread across Wallonia. They are mainly present in the region south of the Meuse and Sambre River. It seems that populations are quite stable these years, and even expanding in some areas[9].

This species is partially protected in Wallonia under regional nature conservation laws. It implies in particular that it is forbidden to hunt, intentionally kill, capture or disturb badgers. However special dispensation can be attributed in some cases. Damages caused by badgers can also be compensated. Indeed, this species is known to cause damages, in particular to crops, which leads to large debates among actors of natural areas' management.

Badgers are known to be regularly flattened by vehicles, especially in spring when young adults disperse. This species has therefore often been at the centre of many studies that investigated impacts of roads on wildlife populations (For example, Dekker et al. (2010), Clarke et al. (1998), Haigh (2012)).

[8] Source : http://biodiversite.wallonie.be/fr/meles-meles.html?IDC=326&IDD=50333780 (30/07/2017)

[9] Source : Schockert V., Lambinet C., Libois R., 2016. *Convention SPW/ULg sur 15 espèces de mammifères protégés ou concernés par la Convention de Berne et sur 2 espèces envahissantes. Rapport d'activités 2016.* Liège : Unité de recherche zoogéographique, Université de Liège. (internal report, unpublished)

1.5.2 / European Polecat[10]

European polecat (*Mustela putorius*) is a smaller mustelid (40 to 60 cm in length; 750 to 1600g. for males and 450 to 800g for females) with a brown-black fur. It can be distinguished from other mustelids thanks to its black and white mask (Figure 5).

Figure 5 : A European polecat. "Mrs Polecat" by Peter Trimming. Published on Flickr under CC BY 2.0 licence.

The polecat can live in various habitat types but appreciate wetlands and valley bottoms most. This is a very discreet species, and regular data encoding or standardized monitoring of this mammal only started few years ago in Wallonia. Only partial knowledge of its distribution and abundance is therefore available at the moment. Nevertheless, it is known that polecats are widespread over the whole Walloon territory[11].

Polecat is a Natura 2000 species, listed in the appendix 5 of the European Habitat Directive 92/43/EEC. It is therefore mandatory to "maintain or restore, at favourable conservation status" this species (Council of the European Communities, 1992). Polecats also have to be monitored, and results have to be communicated every 6 years through European reports.

This mustelid is also considered as a game species in Wallonia, but hunting it is not allowed at the moment. However, in accordance with regional order, special dispensations are regularly attributed to remove individuals that treat farming activities.

According to literature and field observations, polecats are frequently killed by vehicles' traffic. The latter could be an important cause of death for this species (Barrientos et al., 2009; Kristiansen et al., 2007).

[10]Source :http://biodiversite.wallonie.be/fr/putorius-putorius.html?IDC=326&IDD=50333815 (07/08/2017)

[11] Source : Schockert V., Lambinet C., Libois R., 2016. *Convention SPW/ULg sur 15 espèces de mammifères protégés ou concernés par la Convention de Berne et sur 2 espèces envahissantes. Rapport d'activités 2016.* Liège : Unité de recherche zoogéographique, Université de Liège. (internal report, unpublished)

1.5.3 / Raccoon[12]

Figure 6 : A raccoon. "Raccoon on Olympic Peninsula" by Bruce Trimble. Published on Flickr under CC BY-NC-ND 2.0 licence.

Raccoon (*Procyon lotor*) is a grey procyonid, generally 50 to 70cm long and that weight 3,5 to 9kg. It can be easily recognized thanks to its sharp black facial mask and its black and white stripped tail (Figure 6).

Raccoon is not a European indigenous species. It comes from North America and has been introduced in Europe for fur trade. Since then, this mammal has rapidly spread across the continent and is currently colonizing the eastern part of Wallonia (coming from Germany and Luxembourg) and the western part (coming from France). Uncommon in the past, raccoons are nowadays well established in Wallonia and regularly observed for several years.

Like polecats, raccoons appreciate wetlands and valley bottoms most. However, they can adapt themselves to a lot of different environments and situations, even in urban areas. They are also omnivorous and opportunists, adapting their food habits to available resources.

As a result, raccoon is considered as an invasive species in Europe. It is on the list associated to the European regulation n° 1143/2014 that aims at limiting impacts of exotic species. Member states therefore have to take action against raccoon spread, to monitor its expansion and to produce European reports every six years (European Parliament and the Council of the European Union, 2014). In Belgium, raccoons are on the black list of the Belgian Forum on Invasive Species[13]

Mortality due to road has sometimes been investigated in North America (see for example Gehrt (2002) or Rolley et al. (1992)) but has not been researchers' focus in Europe.

---

[12] Source : http://biodiversite.wallonie.be/fr/procyon-lotor.html?IDD=50333996&IDC=326 (08/08/2017)

[13] http://ias.biodiversity.be/species/show/29 (09/08/2017)

As a result, three different kinds of species have been studied in this master thesis:

- Badger is a protected species that is regularly flattened on roads. It sometimes causes damages and worries many actors involved in nature management.
- Polecat is a Natura 2000 species that has to be monitored in order to evaluate populations' status every six years.
- Raccoon is an invasive, fast-spreading species that colonizes new territories. A close watch also have to be kept on this species in accordance with the European regulation.

It was relevant to focus this master thesis on these mammals because of this diversity of situation. Moreover, these species have to be monitored, so it would be interesting to see whether roadkill data could be used for this purpose. They are also regularly flattened, and it would be worth getting more information about how these collisions impact populations.

1.6 / Research questions

All in all, this master thesis attempted to answer the following questions:

a) What methodology should be used to sort out and select relevant mortality data when they have been harvested voluntarily?

b) How to make these data valuable? More precisely :
   - Could roadkill observations serve as an indicator of wildlife populations' status and trends?
   - How do traffic mortality impact wildlife populations?

This thesis' report has therefore been divided into three main chapters that correspond to those three questions. In order to make explanations clearer, these chapters will have their own methods, results and discussions parts.

2 / DESIGNING A METHODOLOGY

2.1 / METHODS

2.1.1 / Data selection from databases

The first step was to extract relevant observations from databases. In observations.be database, collisions with wildlife are normally recorded as "roadkill". Nevertheless, my supervisors and I thought it was too restrictive to only use these data. Indeed, as mentioned before, people encoding on the website have to describe the behaviour of the animal they have seen by choosing among different "activities" in a list. Those include, among others, "road kill" but also "discovered dead" and "present".

We suspected that people could have missed the "roadkill" category and encoded their observations as "discovered dead" or "present" instead. "Discovered dead" could especially contain a significant proportion of observations that actually were traffic victims. "Present" (translated "sur place" in the French version) is the default behaviour proposed in the form. It means that the animal was just there and that it didn't show any particular behaviour. But in French, "sur place" also means "not moving". We thought confusion might arise from this description: people may have believed that "sur place" is the appropriate category as an animal killed is - of course - not moving anymore. We therefore extracted all data from these three categories ("roadkill", "discovered dead" and "present") for the three study species.

The same kind of problem arose when selecting observations from OFFH. In this database's form, there is also a list of different "animal's behaviour" among which observers have to choose when encoding. However, this list doesn't include any roadkill category. The only alternative is then to use "carcass" or "specimen" descriptions. "Specimen", as "present" in observations.be, is a generic category for recording an individual without any particular behaviour. We nevertheless expected that it contained lots of roadkill observations. Thus, all data listed as "carcass" or "specimen" were extracted from the database. OFFH also contained some observations that were known to be related to traffic victims, in particular those gathered during the "Go For Badger" project. They were all taken into account.

All dissections referenced in the Zoogeography Research Unit database were selected as they were all carcasses, often removed from roads.

Only data harvested between 2006 and 2016 (included) have been selected. Observations.be, which is the main data source, started to run in 2006. It was therefore not relevant to conduct analysis before this year.

To simplify explanations and make the following report more fluent, common names have been adopted for data of the same nature. So, "present" and "specimen" will be mentioned as "specimen" in the rest of the text as they correspond to the same kind of information. "Discovered dead" and "carcass" will be named "carcass". And "roadkill" remains "roadkill". To make it clearer, data selected for this study have been summarized in Table 2 here below.

Table 2 : Data sources, data selected and common names used for each type of observation.

Databases	Data selected		
observations.be	"roadkill"	"discovered dead"	"present"
OFFH	traffic victims (Go for Badger project)	"carcass"	"specimen"
Zoogeo. Unit	-	"carcass"	-
	↓	↓	↓
Common names	"roadkill"	"carcass"	"specimen"

To encompass the lack of encoding accuracy in these three databases, I therefore needed to select a wider range of data ("present", "specimen", "carcass", etc.) in order to gather as many potential roadkill observations as possible. This means that I had to design a sorting method to find out what data were actually relevant or not for this study.

This part of the master thesis therefore had two main goals:

a) Designing a standard, efficient and sometimes automatic methodology that could be used in the future for further research and routine monitoring of traffic victims. Having such a methodology would allow 1) to spend less time on data sorting and 2) to compare results year after year as they are based on the same protocol.

b) Applying this method to the different datasets introduced above so that analyses could be performed in the second and third part of this report.

In order to test for the efficiency of the designed methodology, some statistics have been calculated after the different main steps. To see whether automatic processes were accurate and able to select data properly, comparisons between manual and automatic methods have also been done whenever possible.

The different steps will now be explained with more details.

2.1.2 / Checking and merging datasets

Extracting information from databases resulted in a number of various data sheets for each species. It was then necessary to merge them to obtain one single file per species. As there were sometimes a lot of differences between these dataset's contents and layouts, it was very important to take some time to homogenize them carefully.

At this step, all data manipulations have been done using Microsoft Excel 2010 software. Indeed, Excel was the most efficient solution to manage data as it was necessary to visualise them easily and to carry out specific modifications very often.

Here is a list of the main operations that have been done:

- Layout of each dataset has been homogenized. Contents have also been changed to use the same terminology everywhere. Date and location format have been standardized.

- Observations without coordinates or date have been removed as it would have been impossible to use them for further analysis. Unprecise dates (for example "summer 2015") have not been used either.

- Observations located at the scale of municipalities or with a precision higher than 1000 meters were removed. The aim of this study was to analyse roadkill at a fine scale, which wouldn't be possible with this kind of unprecise data. As the latter accounted at the maximum for 2% of datasets, deleting them wouldn't mean missing lots of data.

- Observations that have been refused or low-graded by databases administrators (for example because of species misidentification or because there were some doubts about data credibility) have been removed.

- Observations for which authors have mentioned that they were not sure about species' identification have been removed.

- If number of individuals was missing in observations' description, it has then been assumed that only one animal had been seen.

- Some observations were also sometimes recorded with a number of individuals higher than one. For example, it appeared when the observer saw many dead animals along the same road segment. Instead of encoding a new observation for each carcass, he/she encoded only one point and entered the number of individuals seen in observation's form. It is neither convenient for finding double-counts nor for performing local roadkill analyses. As those observations represented not more than 1% of datasets, they have been deleted.

- Observations that fell outside of Walloon borders have been removed.

The different Excel sheets have finally been merged to produce one global set of observations per species.

Then, the "Find duplicates" query wizard of Microsoft Access 2010 software has been run to search for duplicates. Indeed, a same observation has sometimes been encoded in different databases or many times in the same database by error. Observations' coordinates or identifier have been used as criteria to search for duplicates. Once found, the duplicate that contained the largest amount of information has been kept, the other corresponding duplicate(s) have been removed.

2.1.3 / Management of "specimen" data

As a reminder, it was suspected that some roadkill observations were encoded as "specimen" by mistake. This is not a convenient situation. These data contain various types of information, and simply displaying them on map doesn't allow locating roadkill observations: animals seen along roads could have been dead or alive. Thus, it was needed to find a way to distinguish observations that were actually roadkill and those that were not.

To achieve this goal, badger datasets have been used. Generally speaking, badger's data have been used all along the design of this methodology to test the different steps. Indeed, this species is frequently run over by cars so it represents by far the largest amount of observation available. It was therefore (statistically) meaningful.

In order to verify the content of "specimen" data from observations.be and OFFH, two sub-datasets have here been used: badger's observations gathered between 2011 and 2016 (included), and OFFH observations gathered between 2006 and 2016 (included).

Important note: using those samples arbitrarily was an error. Indeed, since those samples have not been selected randomly among the total dataset, confidence intervals or statistical differences could not be calculated. Indeed, those statistics require random samples. The same kind of error as also been done later when sorting "carcass" data out. This mistake has unfortunately been spotted too late to be changed.

However, choosing arbitrary samples was also sometimes very relevant to explore what was "inside" data. Moreover, the sub-datasets used here generally contained a very large proportion of the total dataset. Results are therefore probably very reliable.

Then, for each of these two datasets, comments written by observers have been read for every "specimen" data. When it was clear that the observation had nothing to do with a collision, it has been removed. However, it has often been impossible to conclude from comments because the latter were not clear or because there were no comments at all. For these data, a second check has been done by Vinciane Schockert (Zoogeography Research Unit): observations have been displayed on a map and sorted on the basis of her knowledge of badger's ecology. When data were located too far from roads, they have been removed. Observers have also been contacted to get more information about their observations.

Thanks to these operations, it has been possible to estimate how many traffic victims have been recorded as "specimen". Statistics have then been calculated to evaluate the proportion of these data compared to the global datasets.

As not more than 13% of "specimen" data from observations.be were actually roadkill it has been decided not to take them into account and to completely remove them (for more details, see results and discussion paragraphs). It has also been done for the rest of badger dataset.

It has been assumed that the same pattern would have been observed for other species and that conclusions drawn from badger's results could be applied to raccoon and European polecat.

“Specimen” data from observations.be have therefore been directly removed from these species’ datasets without neither any look at comments nor contact with observers.

Results were quite different for OFFH “specimen” data as they actually contained an important proportion of roadkill observations. As a consequence, they have not been removed. Instead, it has been decided to check comments related to every “specimen” observation for each of the three studied species. When it was clearly written that animals seen were traffic victims or dead (but not possible to know whether it was a roadkill or not), then data were kept for further analyses. In other cases or when there were no comments available, they were simply deleted. Authors were not contacted to gather more information as it would have taken too much time.

2.1.4 / Management of “carcass” data

As a reminder, “Carcass” data were also extracted from databases as it was suspected that they contained a significant number of observations that would be traffic victims. It was therefore necessary to find a way to distinguish “carcasses” that were actually road-kills and those that were not (animals found dead elsewhere and for other reasons).

Fortunately, it is easier to disentangle these observations than those encoded as “specimen”. Since it is known that all “carcass” data refer to dead animals, a major assumption can be made: all “carcasses” recorded close to roads are actually traffic victims. Supposing it is true, simply displaying these observations on a map and selecting those located next to roads would allow distinguishing road-kills from other types of casualties.

Checking every observation one by one would be too long and not objective. A standardised threshold above which observations could no longer be considered as roadkill has therefore to be determined. This maximum distance to roads could then be used in automatic processes, making data selection much easier and objective.

Nevertheless, determining this distance is not a trivial question. It is known that animals injured by cars don’t necessarily die directly on the tar. They can sometimes crawl few meters (or more) before dying on road verges or in the surrounding landscape (Santos et al., 2011). In those cases, carcasses are not discovered on roads but some distance away.

Moreover, precision of observations’ location depends on map scale at which people point out animals they have seen. Observers sometimes don’t take time to zoom in the map to locate them accurately. The larger the scale, the lower the precision. Thus, it is likely that observations related to traffic casualties are not necessarily pointed exactly on roads.

The threshold used to select “carcass” data that are “close to roads” should therefore encompass these two components (animals’ movement before dying and the lack of precision when encoding).

To calculate this distance to be used, some “carcass” data that were known to actually be road-kills were needed. For that, badger “carcass” observations encoded in OFFH database between 2006 and 2016 have been selected (note that this sub-dataset as already been used above to analyse “specimen” data). These observations had the advantage to contain lots of comments written by authors. Thanks to these comments, it has been easy to identify carcasses data that were traffic victims.

Then, calculating distances to roads of these observations would allow getting an idea of the distance necessary to select 95% of roadkill data recorded as “carcass”. This distance is the threshold needed.

To do so, these data have been displayed in a Geographic Information System (GIS) as a point layer, using ESRI’s ArcMap 10.3.1. Note that this software has been used for every other spatial analysis of this master thesis. A vector layer (designed by the National Geographic Institute of Belgium) that contained all roads and tracks sections of Wallonia has also been used.

“Near” tool available in ArcMap[14] have then been used to calculate the distance of every observation to the nearest tarmac road (dirt roads suitable for motor vehicles have not been taken into account). A cumulative frequencies curve has then been drawn and the distance needed to encompass 95% of observations has been determined.

The same operation has then been done for the 2136 badger observations directly encoded as “roadkill” in observations.be database between 2011 and 2016 (included). These data were used as a reference to compare overall precision of “carcass” and “roadkill” data.

Finally, it has been necessary to test the ability of the calculated threshold to properly find out road-kills recorded as “carcasses” in other datasets. To see whether it was the case, “carcass” badger data recorded in observations.be database between 2011 and 2016 have been selected. Comments have then been read, and only observations described as traffic victims have been kept. Their distances to roads have been calculated. This way, it has been possible to verify whether carcass data that were actually road-kills have correctly been identified by the automatic process.

Results showed that it was efficient. As a consequence, comments have no longer been read and a 55m threshold has been applied (using “spatial selection” tool in ArcMap) to sort every other “carcass” observation in remaining datasets. Again, it has been assumed that conclusions drawn from badger’s results could be applied to raccoon and European polecat as well. Those species’ observations have therefore been processes the same way.

2.1.5 / Double counts

Three main operations have been done so far with data originally selected:

- They have been checked, merged, and sometimes deleted when not relevant, incomplete or not within the scope of this study.
- “Specimen” data have been sorted and often removed.
- “Carcass” observations have been selected based on their distance to roads.

[14] For more information about how Near tool calculates distances, please visit : https://pro.arcgis.com/en/pro-app/tool-reference/analysis/near.htm

As a result, the remaining observations comprised:

- Some OFFH's specimen data considered as road-kills.
- "Carcasses" observations that were located close to roads.
- Data that were directly recorded as roadkill in databases' forms.

All of them were considered as traffic victims. At this step, a dataset only containing relevant observations has therefore been obtained for each species. However, a last operation had to be carried out: finding double-counts.

Issues related to double-counts

Identifying double-counts has probably been the trickiest part of this methodology design. The issue was that a single dead animal could have been recorded by several observers since carcasses were not removed from roads after being detected. Because of this absence of coordinated sampling programme, it was therefore probable that some different observations were actually related to the same animals.

Another challenge is that people generally do not see a same carcass the same day and do not locate it exactly at the same position. As a consequence, these double-counts are unfortunately not "perfect" duplicates: it is not possible to identify them by simply using a kind of "find duplicates" tool or query available in database management software.

Searching for double-counts with the help of a program

Consequently, it has been necessary to design an original program that could compare data to each other and identify double-counts. I didn't write this program myself as it would have taken a long time and as it was not the main goal of this master thesis. Instead, I collaborated with a computer scientist to elaborate it. He designed a Macro code for Excel that is available in the Appendix 2.

This program works like this: the user first has to define a distance and a time interval that will be used as criteria to search for double-counts. Then the program calculates spatial and temporal gaps between every observation thanks to their associated dates and coordinates. When the gap between two or several observations is smaller than criteria defined by the user, then they are identified as potential double-counts. The program finally returns a list with observations and their potential double-counts.

Determining criteria for searching double-counts

However, an issue arose here: what are the appropriate time and distance criteria? These criteria should be large enough to encompass the fact that a same carcass is probably not located exactly at the same date and position by different observers; but they also must be sufficiently restrictive to not overestimate the amount of double-counts.

Unfortunately there were no references to answer this question. The ideal situation would have been to find carcasses along roads, to note down their exact location and persistence, and then to see with what variability observers encode them in databases. But this kind of experiment has not been conducted. It has therefore been necessary to find other ways to estimate appropriate criteria.

Temporal criterion depends on the rate at which carcasses are degraded. This could vary depending on species' size. Some papers have emphasised the importance of taking carcasses persistence into account when monitoring road-kills. In their study "How long do the dead survive on the road?" Santos et al. (2011) surveyed 4447 carcasses and their time persistence. They estimated that dead carnivores (111 to 7300g weigh in that case) had a median duration of nine days (95% interval: 5 to 19 days). In our case, it has been estimated that those results could be applied to polecats and raccoons. A 9-days criterion has therefore been chosen to identify their double-counts. Badgers are larger, so we estimated that a 2 to 3 weeks criterion would be more appropriate.

Distance criterion is rather related to the accuracy with which observers locate carcasses. It is likely that it is not species-specific. Again, there is no reference, and it seems that no scientific paper explored this issue. Consequently, three different operations have been done with badgers' datasets to try to estimate an appropriate distance criterion:

- When encoding, observers have the opportunity to give information about the precision with which they locate their observations. Those precisions have therefore been studied by creating a cumulative frequency curve with all badger's data available.

- To get an idea of the number of double-counts that could be found, the program has been run using all badger's data available and different criteria values that seemed relevant (14 and 21 days and 250, 500, 750 and 1000 meters).

- Finally, all badger's data encoded in 2016 (the latest and probably most comprehensive year of survey) have been selected. Among them, "reliable observers" have been identified by Vinciane Schockert (Zoogeography Research Unit) on the basis of her knowledge of databases and of actors involved in nature management and monitoring. Then, all data encoded by those people for the period 2006-2016 have been selected. The program has been run to find double-counts with the largest possible criteria (21 days and 1000m). A road segment containing many observations (N4 Namur – Marche en Famenne) has then been chosen. There, every group of double-counts have been identified and the maximum distance between two observations of a same group has been measured for each group (see Figure 7 for a better understanding). Distances measured this way have finally been displayed in a cumulative frequency curve.

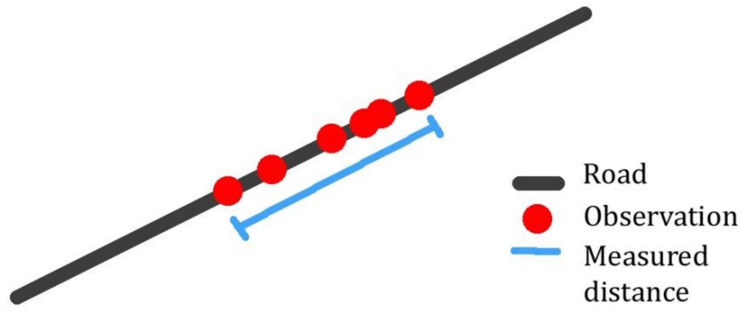


Figure 7 : Scheme of the way that distances have been measured. Points on this road represent observations that have all been identified as double-counts of a same road-kill. The distance between the two extreme observations has been measured to get an idea of the maximum extent of a group of double-counts.

I finally came to the conclusion that there was no accurate objective way to determine an appropriate distance criterion to find double-counts (see results and discussion paragraphs). A 500m- criterion has therefore been determined arbitrarily on the basis of results obtained with the three calculations mentioned above.

All in all, here are the time and distance intervals used for our species:

- Badger: 14 days, 500m
- Polecat: 9 days, 500m
- Raccoon: 9 days, 500m

Selecting and removing double-counts

However, this part of the methodology was not over. Once having identified double-counts, it was still necessary find a proper way to remove some of them so that only one observation would be associated to one roadkill. Indeed, the program only identifies *potential* double-counts. A fine selection has then to be applied.

For example, it could happen that two (or more) animals die within a tiny distance and time gap. The program will detect them as double-counts of a same individual, even if they are actually not. Another issue is that an observation A could be identified as a duplicate of an observation B. B is also a duplicate of C. But C is not a double-count of A. And so on, much more complex situations could occur. How to deal with that? Which observations should be removed, and which should not? It has therefore been necessary to find a way to keep or remove data properly.

Checking every observation one by one on the map would have been too subjective and time-consuming. A standardized and objective method therefore had to be developed. After many attempts, the following solution has been adopted:

- Double-counts are sorted by using the observations' list only. They should not be displayed on a map.
- In any case, observations encoded by the same person are not to be considered as duplicates. It is likely that this person observed several carcasses close to each other. Those observations are therefore not double-counts and have to be kept.

- If only two observations are identified as potential double-counts:
    - The observation that comes from the most reliable source is kept, the other is deleted. The reliability ranking is: Dissection (Ulg) > OFFH > observations.be. The latter database is the least reliable as non-professionals can use it.
    - If the observations come from the same source, the one with the best precision is kept.
    - If they have the same precision, the first observation of the list is kept, the other is deleted.

- If three or more observations are referred to as double-counts:
    - If some of these observations are not identified as duplicates of others (see the example with A, B and C above): some observations are removed so that there is no possible double-count anymore in the result (in our example, A and C would be kept and B would be removed).
    - If all of the observations are identified as duplicates of the others, only one is kept (on the basis of the criteria listed above (source quality, precision, 1st in the list)) and all others are removed.
    - If, within the group of double-counts, several authors each encoded several observations, than observations of the author that encoded the greatest amount of observations are to be kept.

It has also been tried to estimate the efficiency of this method (does it select and remove observations properly?). However, as reality (e.g. actual number, location and persistence of road-kills) was not known, there was no proper reference for comparison. The only solution was therefore to compare the method described above to another one assumed to be more accurate.

This second method consisted of a more manual one. It has only been done for badgers' data in 2016. First, double-counts have been searched using 14 days and 250m as criteria. The selection and removal of double-counts have then been done subjectively by Vinciane Schockert on the basis of her list of reliable observers: when two observations were double-counts, the one with the most reliable observer have been kept and the other has been removed, and so on. Double-counts have then been searched using 14 days and 500m as criteria, and the same kind of selection has been performed again. Finally, results obtained with these two different methods have been compared.

2.2 / RESULTS

Here is the initial number of data extracted from databases, before any manipulation. Observations of each database have been summed up and summarized by observation category (Table 3).

Table 3 : Initial number of data extracted from databases. Observations have been grouped by category.

	Badger	Polecat	Raccoon
Specimen	1239	267	361
Carcass	1011	177	194
Roadkill	2891	469	207
Total	5141	913	762

Badgers accounted by far for the largest amount of available data.

2.2.1 / Tests with "specimen" data

After having merging datasets and removing some irrelevant observations, two badgers' sub-datasets have been selected. Here are the number and proportions of data types they contained (Table 4).

Table 4 : Number and proportions of badger's observations recorded between 2011 and 2016 in observations.be and between 2006 and 2016 in OFFH databases.

	observations.be		OFFH	
	Number	**Proportion (%)**	**Number**	**Proportion (%)**
Specimen	770	24,11	362	45,08
Carcass	288	9,02	234	29,14
Roadkill	2136	66,88	207	25,78
Total	3194	100,00	803	100,00

There were 3194 badger observations recorded between 2011 and 2016 in observations.be database. Among these data, "roadkill" accounted for the largest number. There were 803 badger's observations recorded in OFFH database between 2006 and 2016. They were mainly "specimen" data. Roadkill data correspond to those that were known to be related to traffic victims, in particular those gathered during the "Go For Badger" project.

Comments have then been read (whenever available), and observers have been contacted to determine the actual nature of observations recorded as "specimen". The following categories and figures have finally been obtained (Table 5).

Table 5 : Number and proportions of data types recorded as « specimen » between 2011 and 2016 in observations.be and between 2006 and 2016 in OFFH databases. "Dead" corresponds to animals discovered dead for which there were no indices to determine whether they were traffic victims or not. "Unspecified" corresponds to data for which it has been impossible to determine their true nature, in particular because no answer was obtained from observers. "Other types" groups together observations that were not relevant for this study (for example animals seen alive)

	observations.be		OFFH	
	Number	**Proportion (%)**	**Number**	**Proportion (%)**
Roadkill	54	7,01	118	32,60
Dead	20	2,60	13	3,59
Unspecified	23	2,99	/	/
Other type	673	87,40	231	63,81
Total	770	100,00	362	100,00

Observations.be

Road-kills only accounted for 7,01% of "specimen" data. Assuming that all "Dead" and "Unspecified" observations could be traffic victims, the number of roadkill could increase up to 97 (54 + 20 + 23 observations), which corresponds to 12,60% of the 770 "specimen" data. These 97 potential roadkill are also to be compared with the 2136 certain roadkill of the sample. They equal 4,54 % of these 2136 observations.

OFFH

It appeared that "specimen" actually contained a large proportion of road-kills (32,6 %). There were also some carcasses encoded as specimen.

2.2.2 / Tests with "carcass" data

Badger's "carcass" observations encoded in OFFH database between 2006 and 2016 and in observations.be database between 2011 and 2016 have been used for calculations. Number and proportions of "carcass" data compared to other categories of observations have already been showed in Table 4 in the above section. Comments have then been read to try to determine the exact nature of each "carcass" observation. Results are presented in Table 6.

Table 6 : Number and proportions (based on observers' comments) of observations types recorded as "carcass" in two sub-datasets. "Roadkill" corresponds to "carcasses" that were actually traffic victims; "Undescribed death" are data for which comments didn't allow guessing the cause of the death; "Other causes" corresponds to animals killed by other causes than traffic.

	OFFH (2006-2016)		observations.be (2011-2016)	
	Number	**Proportion (%)**	**Number**	**Proportion (%)**
Roadkill	159	67,95	89	30,90
Undescribed death	68	29,06	184	63,89
Other causes	7	2,99	15	5,21
Total	234	100,00	288	100,00

As expected, the proportion of observations encoded as “carcass” that were actually related to traffic casualties was pretty high. Those percentages could have been even higher since it was not possible to determine the cause of the death (because comments where not clear or because there were no comments) for respectively 29,06 % and 63,89% of the sub-datasets. Those “undescribed death” could also contain roadkill observations.

Then, distances to roads have been calculated for these “carcasses” that were known to be traffic victims. It has also been calculated for the 2136 badger’s roadkill observations from observations.be (2011-2016). Some statistics are shown in Table 7. Cumulative frequencies curves are then presented in Figure 8.

Table 7 : Different distance parameters calculated for the three used sub-datasets (namely badger’s “carcass” observations known to be traffic victims encoded in OFFH database between 2006 and 2016; in observations.be database between 2011 and 2016; and badger roadkill observations from observations.be encoded between 2011 and 2016). Distances necessary to encompass 80%, 95% and 99% of observations are shown at the bottom of the table. Data percentages corresponding to a 55m threshold are finally presented.

	Carcass (OFFH)	Carcass (obs.be)	Roadkill (obs.be)
N observations	159	89	2136
Median distance to road (m)	6,44	9,24	5,55
Min. distance to road (m)	0,03	0,14	0,02
Max. distance to road (m)	339,17	183,02	277,45
Distance to 80% obs.	21,04	32,63	13,04
Distance to 95% obs.	55,27	98,22	44,01
Distance to 99% obs.	280,11	138,21	150,20
% data corresponding to a 55m threshold	0,95	88,76	96,28

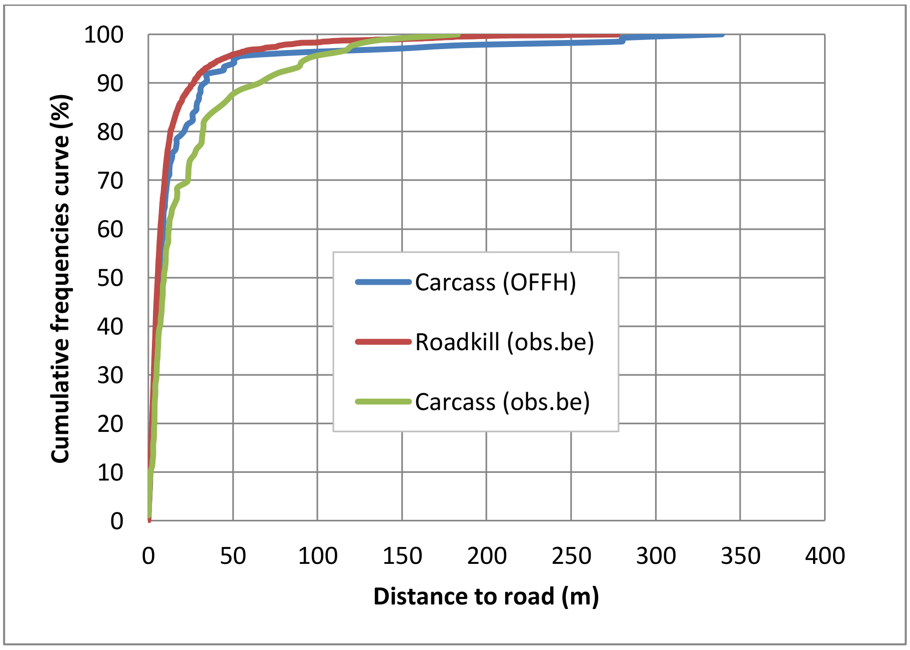


Figure 8 : Cumulative frequencies distributions of the three used sub-datasets (namely badger “carcass” observations known to be traffic victims encoded in OFFH database between 2006 and 2016; in observations.be database between 2011 and 2016; and badger roadkill observations from observations.be encoded between 2011 and 2016).

As suspected, “carcass” data actually related to traffic victims were not necessarily located exactly on roads but some distance away. Even observations directly recorded as “roadkill” in observations.be were not always mapped exactly on roads. Some observations were sometimes pointed out more than 200m or 300m away from any bitumen surface. Nevertheless, as it can be observed on the different cumulative frequencies curves, most of them were located quite close to roads. Median distances of each sample never exceeded 10m.

On the basis of the curves, the threshold to encompass 95% of OFFH “carcass” observations (which is the sub-dataset used as a reference to determine this threshold) has been set to 55m. This distance has then been applied to the 89 “carcass” data from observations.be that were known to be road-kills in order to test for the efficiency of this threshold. 88,76% of them have been selected (See table 7 above). It means that using a 55m distance allowed to correctly select 88,76% of carcass data that were traffic victims. A 98,22m distance would have been necessary to catch 95% of them.

2.2.3 / Double-counts

Here are the results of the three operations that have been done in order to determine what would be the appropriate distance criteria.

3705 badger’s roadkill observations have been obtained after having sorting “specimen” and “carcass” data (see Table 10 below). Among them, 26,23% had an undetermined precision. The remaining 73,77% that had an indicated numerical precision have been displayed in Figure 9 here bellow.

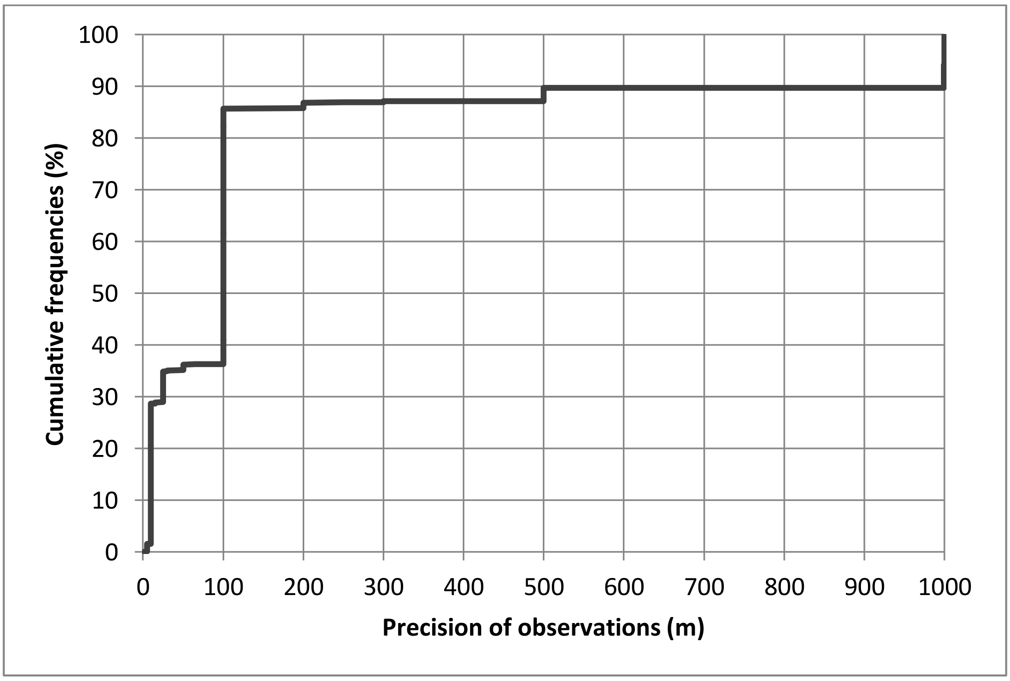


Figure 9 : Cumulative frequency curve of the precisions of badgers’ road-kill observations (all data available together)

85,69% of the observations had a precision inferior or equal to 100m, and 89,64% had a precision inferior or equal to 500m. Nevertheless, those figures should be used carefully (see “discussion” section).

Lengths of groups of double-counts

There were 70 “reliable observers” identified in 2016. Between 2006 and 2016, these authors accounted for 1101 badgers roadkill observations (29,73% of the total dataset). Among those observations, double-counts located between Namur and Marche-en-Famenne have been analysed. 31 groups of double-counts have been found there. Here are the measured distances for each group (Figure 10).

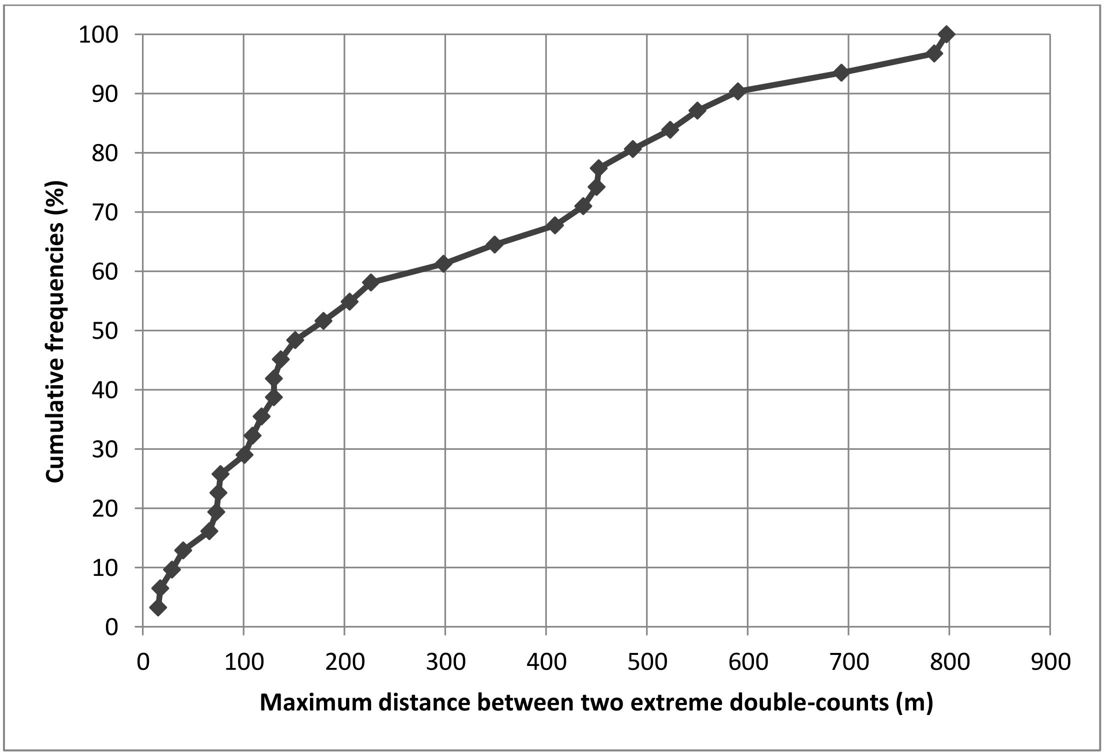


Figure 10 : Maximum measured distances between observations located at the extremities of a group of double-counts. All Badgers’ data from “reliable observers” that are located on the N4 road segment between Namur and Marche-en-Famenne.

The majority of the double-count-groups (54,8%) were less than 200m long. 80,64% of them were less than 500m long. What does that mean? For instance, that it would be worth using a distance of at least 200m to search for double-counts, as the majority of them are located less than 200m from each other.

Number of double-counts depending on criteria

Finally, the program has then been run using all badger’s data available (n = 3705) and different values of criteria that seemed relevant. Results are shown in Table 8.

Table 6 : Number of potential double-counts found using different temporal and spatial criteria. Proportions of the total dataset are also displayed. Those double-counts have been searched among all badger's data available (total = 3705).

Distance (m)	Number		Proportion of total dataset (%)	
	14 days	21 days	14 days	21 days
250	715	752	19,30	20,30
500	960	1018	25,91	27,48
750	1092	1168	29,47	31,52
1000	1212	1294	32,71	34,93

Depending on criteria, the proportion of potential double-counts accounted for 19,30% to 34,93% of the total dataset. Changing the distance criteria led to rapid variations in the number of double-counts, whereas time criteria had less effect on those numbers.

Comparison between two methodologies

Table 9 shows some figures obtained after having sorting double-counts of 2016 using two different methods (the manual and subjective one and the standardized and objective one).

Table 7 : Comparison of two different methods used to sort out potential double-counts identified within badger's dataset (all available observations in 2016).

Numbers	Manual Method (reference)	Standardized method
Observations in 2016	680	680
Potential double-counts	179	179
Double-counts removed	94	89
Double-counts removed (% of 2016)	13,82	13,08
Final number of data	586	591
OFFH data in remaining dataset	117	115
Obs.be data in remaining dataset	350	358
Dissections data in remaining dataset	119	118

A striking similarity between the results can be observed. The amounts of removed double-counts and the final numbers of observations are almost the same. Databases' proportions in remaining datasets are almost equal as well.

Final datasets

Finally, after having sorting double-counts with the standardized method, proper datasets have been obtained. A comparison of the amount of data before and after removing of double-counts is presented in Table 10:

Table 8 : Number of observations before and after having sorting double-counts. Data are displayed by species and databases.

Database	Before sorting double-counts			Final datasets		
	Badger	Polecat	Raccoon	Badger	Polecat	Raccoon
observations.be	2869	458	217	2466	430	189
OFFH	486	86	33	408	46	31
Dissections	350	43	72	344	43	72
Total	3705	587	322	3218	519	292

Badgers finally account for much more data that the two other species. 13,14%, 11,58% and 9,31% of the observations have been removed when sorting double-counts of badgers, polecats and raccoons respectively. Observations.be remains the most important source of data in the end.

2.3 / DISCUSSION

2.3.1 / Tests with "specimen" data

Including "specimen" observations into datasets used to monitor roadkill increases data preparation complexity. To identify observations that actually are traffic casualties, it is necessary to check every observation one by one and sometimes to contact observers. As a consequence, sorting these data out is challenging and time-consuming. It would be worth only if they would contain a significant proportion of roadkill observations.

Observations.be

For observations.be, it has been shown that it is not the case. Traffic victims accounted at the maximum for 12,60% of "specimen" data. It also corresponded to a very little proportion of observations directly encoded as roadkill (4,54 %).

I therefore recommend not to take "specimen" data from observations.be into account when working on roadkill-related topics. It means that some traffic victims' data would be missed, but only a very small proportion. Finding those victims data recorded as "specimen" is like looking for needles in a haystack. Avoiding this step would be easier, more automatic and more objective, which is one of the aims of this methodology.

OFFH

Conclusions were different for "specimen" data from OFFH database. Indeed, many of them were actually roadkill observations. Not using "specimen" data would mean going without a high proportion of traffic victims contained in OFFH database. It would therefore make sense to use and sort these "specimen" data.

Nevertheless, one could say that the absolute number of roadkill observations recorded as "specimen" is almost the same for both batabases (97 at the maximum in observations.be and 118 in OFFH). So why not processing data the same way? Because if only OFFH data are used to monitor roadkill, then it gets completely relevant to use and sort "specimen" data as they contain lots of traffic victims.

All in all, I recommend using "specimen" data from OFFH. As there are not too many observations encoded every year, it could be possible to check every comment and to contact observers year after year if necessary.

On top of that I strongly recommend improving data encoding in the future as it is at the root of those problems. Roadkill observations should no longer be encoded as "specimen". The best case would be a situation in which this kind of mistake doesn't appear at all, so that researchers wouldn't have to take care about those "specimen" data anymore. It would enable avoiding wasting lots of time and energy.

Moreover, sorting "specimen" data (namely by reading comments and contacting authors) is only feasible as long as there are not too many observations to be checked. If datasets go larger, than sorting them could get too complicated.

Some efforts should therefore be done to make users aware of this issue in order to improve observations' encoding. At least, road-kills should be recorded as "carcasses", which are easier to sort out. But the ideal solution would be to add a "roadkill" category in OFFH's form.

2.3.2 / "Carcass" data

"Carcass" data accounted for a non-negligible proportion of total available observations. Moreover, reading their comments and determining their true nature showed that a lot of them were actually roadkill (at least 67,95% and 30,90% respectively for the two used samples).

In absolute numbers, these data didn't represent a large amount compared to total datasets. Nevertheless, they could be easy to sort (contrary to "specimen" data) thanks to a spatial selection. Taking "carcass" data into account for this study was therefore definitely relevant.

Thanks to OFFH sub-dataset, it has been determined that a distance of 55m would enable to select 95% of carcass data that were road-kills (and to reject others). This is of course an estimate based on a sample. However, this is the most accurate and standardized figure available.

When applying this distance of 55m on another sub-dataset, it enabled to properly select 88,76% of "carcass" data that were actually traffic victims. Thus, the test was quite successful (even if confidence intervals or significant tests should have been performed to prove it). It has therefore been decided not to read comments anymore and to use a 55m-threshold to select or reject "carcass" observations in the remaining datasets.

This automatic method assumed that all “carcasses” found along roads were traffic casualties. This was a major assumption but it could reasonably be accepted. It would have been meaningful to test whether there were animals dead for other reasons than collisions in roads’ surroundings. But the lack of this kind of data didn’t enable to carry out any reliable test.

It could have been possible to use a threshold higher than 55m in order to select more than 95% of relevant observations. But this would have radically increased the distance (see cumulative curves) and the risk to select carcasses that had nothing to do with road casualties. Some relevant data were probably missed when using a 95% threshold, but they were those with the worst precision. It was therefore not a big deal.

In the Flemish project “Dieren Onder de Wielen”, neither “specimen” nor “carcass” observations have been taken into account when preparing data (Vercayie et al., 2017; Vercayie et al., 2012). Results described above suggest that it would have been useless to analyse “specimen” data but that “carcass” observations could have been sorted easily. This would have increased the number of available data.

Again, it is to be noticed that data quality and encoding are at the root of all those issues and discussions. Even if sorting “carcass” data out is not so hard thanks to GIS and spatial selections, not having to take care of them would be even easier. Roadkill observations should therefore no longer be encoded as “carcasses”.

Observations.be already contains a “roadkill” category. Database’s users should therefore be sensitized to improve data quality. On the contrary, there is no “roadkill” category in OFFH’s form, so that users are forced to record their roadkill observation as “carcass” or “specimen”. Thus, it was not astonishing to see that “carcasses” contained a very high percentage of road casualties. Adding a “roadkill” category to OFFH’s form would be a very easy solution to solve this issue.

2.3.3 / Double-counts

Determining appropriate searching criteria

When looking at the precision graph, it seemed that most of the observations had a good precision, often inferior equal to 100m. Thus, taking a 100m-distance as criterion would be a reasonable solution. Nevertheless this graph should be used carefully.

Indeed, the large amount of data with a 100m precision value (that accounted for 49,40% of the dataset) is questionable. Actually, 9,19%% of them (GoForBlaireau observations) were encoded with a 100m value by default (Alain Licoppe, pers. disscuss.), and 87,56% of them were encoded on observations.be. On this website, a precision value is proposed automatically to the observer in the encoding interface, depending on the map scale at which he/she ticks off his/her observation on the map. This proposed value is often 100m. It is likely that some observers didn't take care of the precision while encoding, thus increasing the number of observations with a 100m precision. However, it is impossible to find out how many observers did this mistake and what the actual precisions of their observations are.

Considering this bias and the amount of data that had an undetermined precision (23,26% of the global badger's dataset), it was not possible to conclude only on the basis of precisions data.

Cumulative frequencies curve of the lengths of double-counts groups gave another kind of information. It can be observed there that the majority of the double-counts groups had a length inferior to 200m. Some of them were 800 or 900m wide. This is valuable information. However, one should be keep in mind two important arguments:

- Only "reliable observers" data have been used to measure these distances. It is likely that "less reliable observers" encoded their observations with a poorer precision. Thus, the graph of cumulative frequencies could be different when taking all available observers into account. It has been tried to measure those distances with all observations together, but the situations were often too complex to clearly identify groups of double-counts and to measure distances.

- This graph directly depends on criteria used to search for double-counts (here 1000m and 21 days). If other criteria had been used (for example 500m and 14 days or 2000m and 21 days), then other groups of double-counts would have been found by the program, other distances would have been measured and the resulting graph would have been different. Those results are therefore only partially reliable.

Finally, calculating the amount of double-counts found depending on criteria values enabled getting an idea of the number of double-counts, but not more.

As a consequence, no clear answer has been found to know what the appropriate distance criterion was. I came to the conclusion that it would be necessary to define it arbitrarily. Using a long distance would increase the number of double-counts encompassed, but also the sorting complexity and the risk of making errors. A kind of balance had to be found between accuracy and energy spent to sort out double-counts. Based upon results presented above, it has been estimated that 500m was a reasonable solution. It is acknowledged that some double-counts were probably missed when using this distance. However, there was no mean for evaluating that.

Selecting and removing double-counts

It is also acknowledged that the methodology designed here to select and remove double-counts is not perfect. It tries to be standardized, objective and accurate in the meantime. But unfortunately, reality will ever remain unknown: actual number, location and persistence of road-kills are uncertain. Sorting has to be done without any reference, and you can only guess when selecting or removing data. Thus, it is likely that there is no perfect method.

Comparison with the manual method that served as a reference showed very good results. Final figures were almost the same. This similarity is maybe due to the fact that "reliable observers" often encoded their observations in OFFH and Dissection databases, which have also been considered as more reliable sources in the standardized method. Even if it is hard to determine whether this similarity between the two methods is a coincidence or not, it has been accepted that the standardized methodology designed here was efficient.

In their project "Dieren onder de Wielen", Vercayie et al. (2012 and 2015) analysed roadkill data recorded in observations.be in Flanders. They estimated that double-counts accounted for 4% of datasets. Nevertheless, they didn't attempt to remove them. Results of this master thesis suggest that this proportion is probably higher, as the number of double-counts finally removed ranged from 9% to 13% of datasets depending on the species.

To improve observations' precision (which would reduce uncertainties when sorting double-counts), it is suggested that drivers could use smartphone apps to directly register roadkill observation when driving, instead of encoding at home. These apps work with GPS and enable a quick and safe encoding while driving. Their efficiency has been tested and approved by Olson et al., (2014) in their paper "Monitoring Wildlife-Vehicle Collisions in the Information Age: How Smartphones Can Improve Data Collection".

2.4 / CONCLUSIONS OF THE METHODOLOGY

In this chapter, a standard and often automatic methodology has been designed to properly sort out datasets and to select relevant observations for studying road-kills. Key results and discussions are summarized here.

To make a long story short, it is recommend using the following methodology to sort out data:

- Extract all "carcass" and "roadkill" data from observations.be. Do not take "specimen" data from this database into account. Extract "specimen" and "carcass" observations (+those from the Go For Badger project) from OFFH database. Use all dissection data of the Zoogeography Research Unit.

- Check and merge datasets. Delete observations when they are not relevant or incomplete (see the complete list in paragraph 2.1.2).

- Use "specimen" data of OFFH database, read authors' comments whenever available, and keep for further analyses observations that are clearly described as dead animal or road-kills. Remove others. Authors may also be contacted.

- To find out "carcass" observations that are actually traffic victims, use a spatial selection. Those located less than 55m away from a road should be kept. Others are to be removed.

- To identify double-counts, run the designed Macro code in Excel. Determine temporal criterion using the paper of Santos et al. (2011). Use a 500m spatial criterion. To remove double-counts properly, use the method described in paragraph 2.1.5. As a result, a dataset is obtained that can now be used for analyses.

One of the major conclusions of this methodology is also that data encoding should be improved so that traffic victims are neither recorded as "specimen" nor as "carcass" anymore. This way, it would not be necessary to sort those categories out. Only "roadkill" data could then be directly used. For this purpose, observers should be sensitized and a "roadkill" category should be added to OFFH database.

It would also be useful to improve precisions of observations' location as much as possible. Applications available nowadays for smartphones could be part of the solution.

On top of that, it would be even more efficient to create a single database to record all biological observations. At the moment, biodiversity's data are scattered among many different actors and institutions. It leads to problems when studies are conducted: only partial information is sometimes available, and time is wasted to merge data that have been encoded using various forms and protocols. It would be much more efficient and wiser to create a common platform. At least, forms of each database should be homogenized and standardised (e.g. using the same fields, attributes, layout, etc.), so that merging datasets from different sources would be easier.

As far as I know, existing studies using voluntary observations didn't explore data quality as in depth as what has been done here. In particular, it is not common to take into account the fact that some road-kills could have been encoded many times. Studies generally use full datasets without attempting to precisely locate double-counts. The methodology designed here is therefore quite original.

3 / ROAD-KILLS AS INDICATORS OF WILDLIFE POPULATIONS

3.1 / INTRODUCTION: DEALING WITH DATA HARVESTED VOLUNTARILY

3.1.1 / Potential biases and errors.

Analysing data gathered on a voluntary-basis implies dealing with some biases and uncertainties, especially when observations are recorded by non-professionals. Some of them have already been cited in past chapters. Here is a non-exhaustive list:

- Wrong species determination
- Observations misplaced or with a poor precision (for example because it is sometimes hard to find out your exact location while driving or because of the use of an inappropriate map scale when encoding)
- Wrong description of the state of the animal (road-kills are encoded as “specimen” or “carcasses”)
- Double-counts: as dead animals are not removed from the road after being detected, a same individual could have been observed and encoded by different people.
- A greater amount of traffic increases chances to observe carcasses before they disappear. It is therefore probable that detectability is much higher on major roads than on smaller ones.
- Recording road-kills or not depends on observers’ mood and sensibility to nature-related topics.

Moreover, some errors are inherent to the harvest of roadkill data. They do not depend so much on voluntary sampling:

- After a collision, animals not necessarily die on the tarmac. Sometimes they escape in the surrounding landscape without being detected.
- Carcasses are degraded by vehicles and/or consumed by scavengers. Thus, they could disappear ore become unrecognizable before any naturalist record them (Santos et al., 2011). The same problem occur when drivers collect carcasses for meat consumption or when road services remove dead animals that could affect traffic safety (road administration, pers. discuss.). All those facts lead to an underestimation of the number of carcasses.
- Carcasses’ detectability increases with animals’ size: it is then easier to notice them and to identify the species. As a result, it is likely that large animals are over-represented in databases compared to smaller ones.
- Vehicles’ speed decreases observation easiness: the faster a car is, the less time to count road-kills and determine the species.
- And so on, and so on…

As a consequence, lots of uncertainties are associated to roadkill data used for this master thesis. Unfortunately, some of them are unavoidable. And in any case, these issues cannot be solved *a posteriori* once data harvest is over. The only thing that can be done is to be aware of those potential errors and to analyse data carefully.

However, the most problematic issue is probably the absence of coordinated sampling program. It has already been said that in most of road ecology studies, data are harvested following a rigorous protocol (for example: Clevenger et al., 2003; Grilo et al., 2009; Haigh, 2012; etc.) This was not the case here. It means that search effort was not homogenous on the whole territory. As a consequence, a high concentration of road-kills on a certain place could be explained by a local high density of animal populations and/or by a local high search effort and/or by road infrastructures that are very risky for wildlife to cross. Unfortunately, it is tricky to disentangle these elements without having any idea of sampling effort.

3.1.2 / Conceptualizing the issue

To get a better understanding of the situation, the issue of wildlife-vehicles collisions have been conceptualized in cooperation with Marc Dufrêne (Gembloux Agro-Bio Tech, University of Liège). To do so, relationships between four key-components (road-kills, wildlife populations, vehicles and road infrastructures) have been identified (Figure 11).

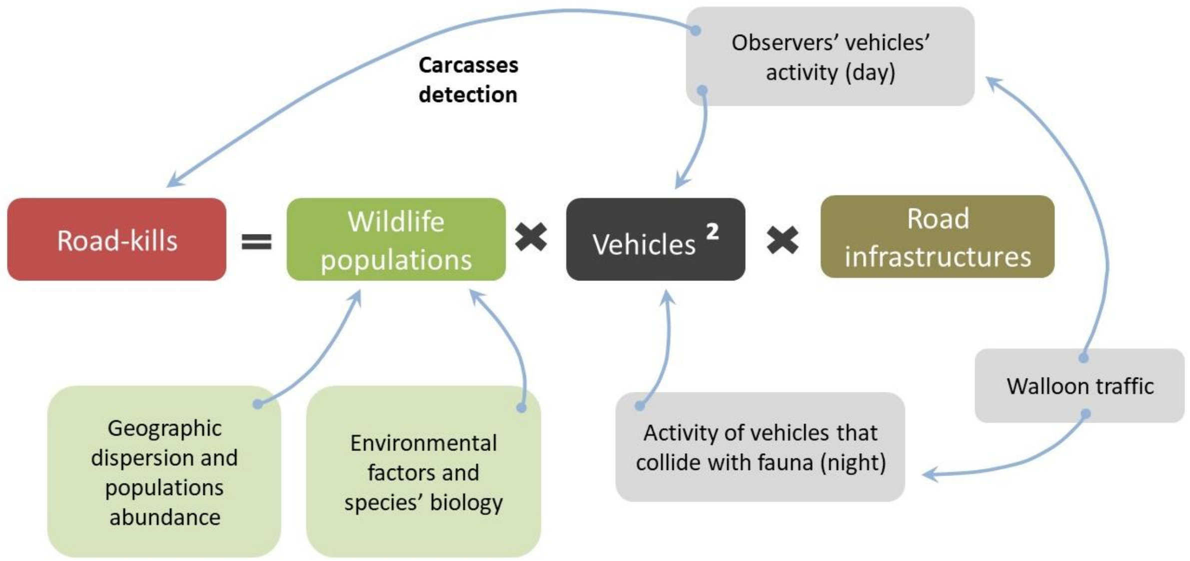


Figure 11: A conceptualization of wildlife-vehicles collisions issue. Figure design: Henry Ziegler

The occurrence of a collision depends on the interaction of three main variables: wildlife populations, vehicles and road infrastructure. If there is no fauna population where roads and vehicles are, then collisions could not occur. The opposite is true as well. These three variables vary in time and space, which explains that roadkill occurrences also vary in time and space.

“Population” variable depends on two sets of factors:

- Species’ abundance and geographic dispersion (collision risks increase if animals live close to roads and if they are numerous)
- Environmental factors and species’ biology: preferred habitat, dispersion, behaviour against roads, etc.

“Road infrastructures” variable depend on the width of the road, the presence of fences, of fauna passages, of berms, etc.

“Vehicle” variable is directly related to road traffic in Wallonia. The latter varies depending on the place and the moment of the day and of the year (for example traffic is higher on major roads and is different during the day and the night). “Vehicle” is actually a squared variable as it encompasses two important components:

- Cars that collide with wildlife. It is likely that it corresponds to traffic at dusk, night and dawn as wild animals are usually more active at this time.
- Vehicles of the drivers that observe road-kills. It probably corresponds to diurnal traffic, when visibility is high. Those are the drivers that detect carcasses and that finally record them in databases. Thus, they determine the amount of available roadkill data.

The best option would have been to analyse roadkill data thanks to this model. This would have partly solved the problem of unknown sampling effort. It also would have required modelling calculations to take lots of information into account. I didn’t perform such calculations, because I have been lacking time and because I only managed to obtain traffic data at the very end of my master thesis.

I therefore only performed descriptive statistics, acknowledging the limits of the results. Moreover, I had to use available roadkill data as “presence” data only. It means that if a collision was recorded at a place, than the presence of a roadkill there is sure. On the opposite, absences of collision data could be due to a “true” absence or to a low sampling effort. No conclusion should therefore be drawn from those areas without data.

3.2 /METHODS

3.2.1 / Maps and numbers of roadkill data

Firstly, all badgers’, polecats’ and raccoons’ roadkill data that have been obtained at the end of the sorting methodology (see chapter 2) have been displayed on a map. A bar chart with the annual number of observations has also been produced for each species.

It has been suspected that those numbers of observations could actually be related to the annual amounts of observers that record them (this number of observers would roughly represent search effort). To see whether there was a relationship or not, linear correlations have been calculated between those two variables. Calculations have been run for the period 2006-2016 (all available years) and 2009-2016. Indeed it has been determined that the latter period was probably more relevant to analyse data as only very few observations have been encoded in databases before 2009.

3.2.2 / Traffic victims as indicators

As already said, roadkill data could serve as an indicator of living populations only if a clear relationship between those data and actual populations' status exists (Baker et al., 2004). A good temporal and spatial similarity is therefore needed.

However, wild mammals are elusive and not easy to observe. It is therefore tricky to obtain hard facts about their abundance and dispersion (George et al., 2011). Even in the case of red deer which is a famous game species for which lots of censuses are conducted, populations' parameters often remain uncertain (Alain Licoppe, pers. discuss.). Monitoring of mustelids only started few years ago. Knowledge about their populations is still incomplete[15]. There is therefore no clear reference for roadkill data to be compared with.

Thus, two different approaches have been attempted: comparisons with all data available for polecat and raccoon, and a comparison with a cartographic model in badger's case.

3.2.3 / Analyses for polecats and raccoons

All data encoded in OFFH database, observations.be and the database of the University of Liège have been extracted. They have served as a reference for comparison. However, those data also have their own uncertainties as they were often recorded on a voluntary basis as well, without any sampling method. Thus, they suffer the same potential biases than roadkill observations.

To prepare datasets, some data have first been removed, namely those located out of Walloon borders, those refused by administrators, and those with an uncertain species determination. Observations directly recorded as "roadkill", and all "carcass" and "specimen" data that have been considered as road-kills during the methodology design, have also been removed. Observations of more than one individual have not been deleted (on the contrary of traffic victims data). When there was no number indicated, it has been assumed that only one individual has been observed. Finally, duplicates within and between databases have been removed and only data recorded between 2006 and 2016 (included) have been kept.

As a result, proper datasets containing all available observations (excepted those previously selected as road-kills in the last chapter) have been obtained. These observations will be referred to as "other types" of data in the following paragraphs.

[15] Source : Schockert V., Lambinet C., Libois R., 2016. *Convention SPW/ULg sur 15 espèces de mammifères protégés ou concernés par la Convention de Berne et sur 2 espèces envahissantes. Rapport d'activités 2016.* Liège : Unité de recherche zoogéographique, Université de Liège. (internal report, unpublished)

To see whether those “other types” of observations were also directly linked to the number of observers that recorded them, linear correlations have been calculated between the annual numbers of observations and the annual numbers of authors.

Temporal similarity

The temporal similarity between roadkill data and other types of data has been tested through linear correlations between the annual numbers of observations of each category. Correlations have been calculated using

- 2006-2016 and 2009-2016 periods;
- Data summed over six years successively (2006-2011; 2007-2012;… ; 2011-2016) as it corresponds to the time-gap of the European reports.

Since there was a strong correlation between the number of observations and the number of observers for roadkill data and other types of data at the same time (see results’ section), it has been decided to calculate indexes:

Index = number of observations / number of observers

This index has been calculated for each year and for roadkill and other types of data. Then, linear correlations have been calculated again.

Spatial similarity

To test the spatial similarity between roadkill data and other data, an UTM grid with 10 x 10km cells has been used. This size of cells corresponds to the one used in European reports. Thus, it was the most relevant option.

For each cell, the number of roadkill and “other types” observations that fell into the cell each year has been calculated with ArcGIS thanks to “spatial join” functions. Numbers have then been converted into presence/non-detection data (1 or 0).

The next step has been to check, for each cell and each year, if there were roadkill data and other types of data within the cell at the same time, or if only one category has been observed in the cell. This comparison has be done automatically using a combination of “IF” functions in Excel.

For each year, I then quantified the proportion of cells that contained roadkill data and other data at the same time, the proportion of cells that only contained roadkill data and the proportion of cells that only contained other types of data. Cells that didn’t contain any observation have not been taken into account to calculate those proportions. Average proportions have then been calculated over the period 2009-2016.

Finally, the same statistics have been performed after having summing presence data over 6 years successively (2009-2014; 2010-2015 and 2011-2016). Again, the goal here was to match the context of European reports. Average proportions have then been calculated over these 3 sets of years.

3.2.4 / Analyses for badger

In badger's case, it has been decided to compare collisions data with a cartographic model. This model has been designed at the University of Liège, Belgium, in 2010[16]. This is a map that represents suitability of the habitat for badgers in the southern part of Wallonia (Figure 12). It has been build using different variables that influence the occurrence of badger's burrows (vegetation cover, slope, hydrography, proportions of forests and agricultural lands in the landscape, etc.). To make a long story short, cells of the model range from a value of zero to twenty (integer values). The lower the value of a cell, the higher suitability it has for badgers' establishment.

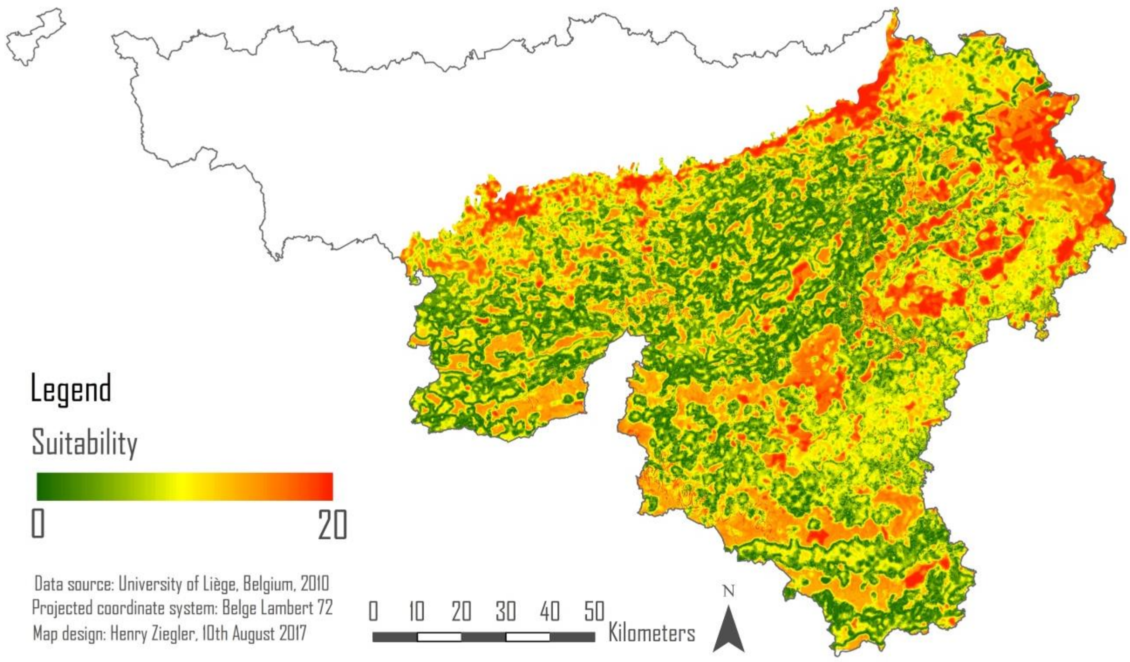


Figure 12 : Model of habitat suitability for badgers in the southern part of Wallonia.

The aim here was to test whether collisions occurred in areas that corresponded to badger's habitat preferences.

[16] Source : Schockert V., Lambinet C., Libois R., 2010. *Convention SPW/ULg sur 15 espèces de mammifères protégés ou concernés par la Convention de Berne et sur 2 espèces envahissantes. Rapport d'activités 2010.* Liège : Unité de recherche zoogéographique, Université de Liège. (internal report, unpublished)

A roadkill observation could be located at a position with a certain value that greatly differs from the direct surrounding values of the model. Only taking into account the exact value at point's location would be an error. Thus, it was necessary to obtain the value of the model in the surroundings of each observation point.

To do so, the model has been summarized using a 1 x 1 UTM grid. For each cell, the median value of the model has been calculated. One squared kilometre almost corresponds to badgers' average home range (110 ha). It was therefore relevant to use this cell size. Then, using spatial joins in ArcGIS, the value of a cell (= the median value of the model) has been attributed to each observations that fell within this cell. Doing so, an estimation of habitat suitability in the surroundings of each traffic victim has been obtained.

Finally, a chart has been produced with the number of collision victims for each value of the model. As a reminder, roadkill data used here can only be interpreted as "presence" data. Thus, only the cells that actually contained observations have been used to produce those statistics. All badger's roadkill observations available between 2006 and 2016 have been taken into account.

3.3 / RESULTS

3.3.1 / Maps of roadkill observations

First, here are the maps of roadkill observations obtained for each species after having sorting data and double-counts (Figure 13, Figure 14 and Figure 15)

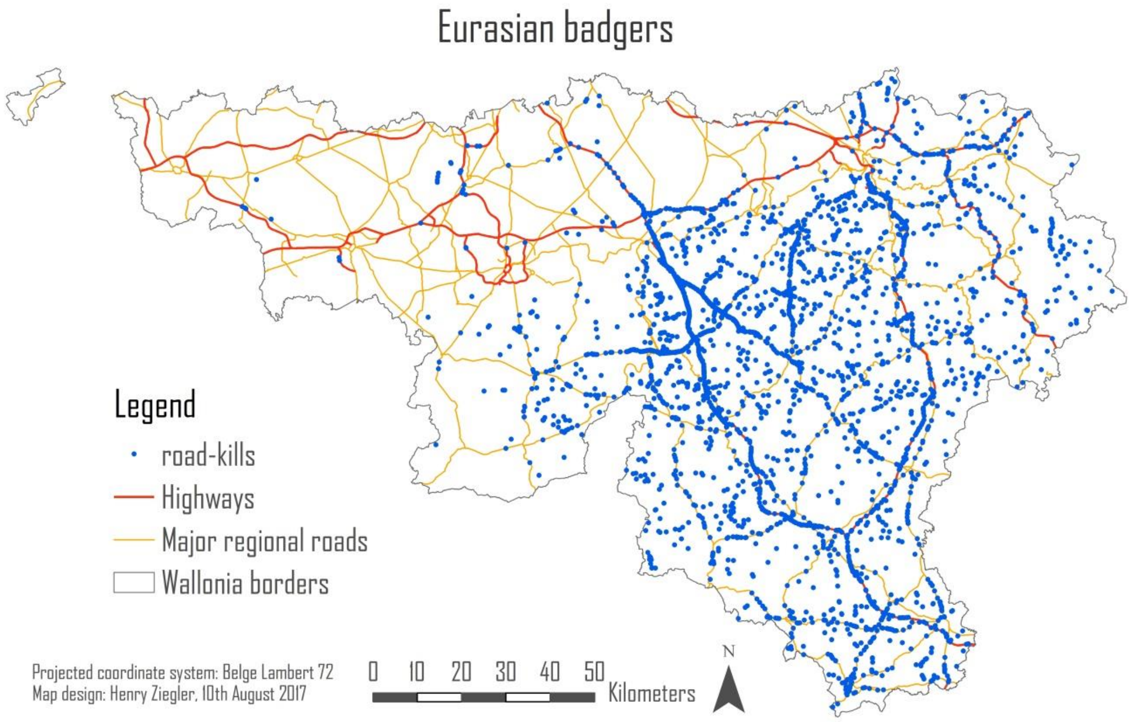


Figure 13 : Map of badgers' road-kills observations for the period 2006 - 2016

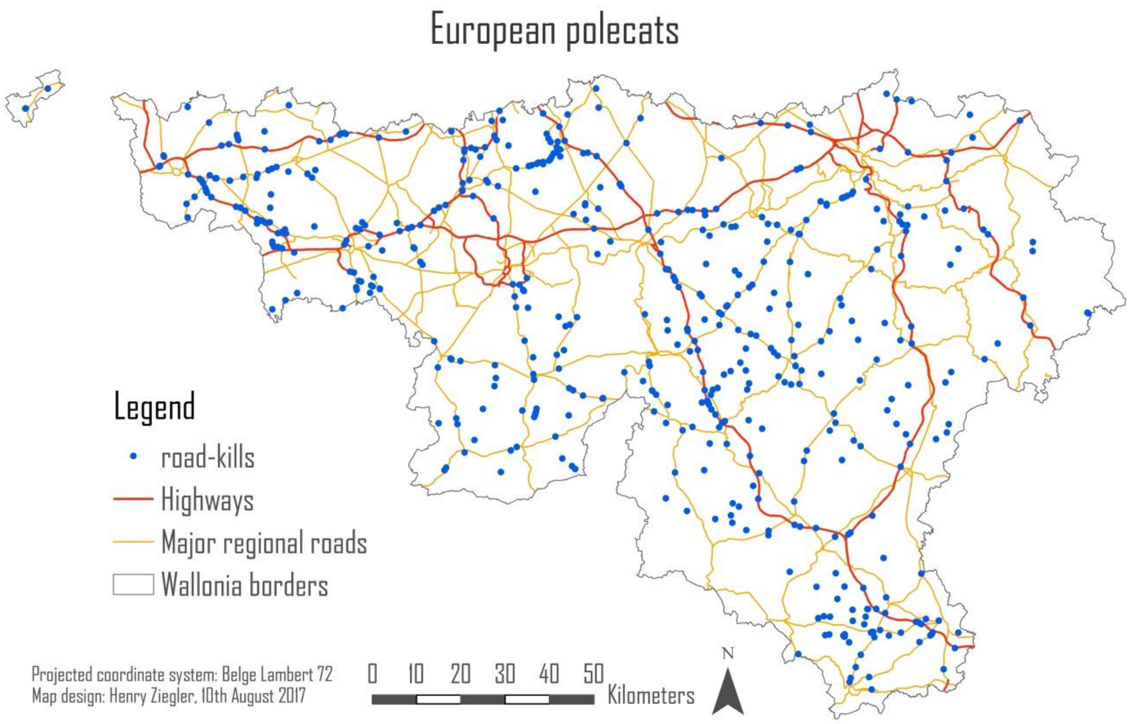


Figure 14 : Map of polecats' road-kills observations for the period 2006 – 2016

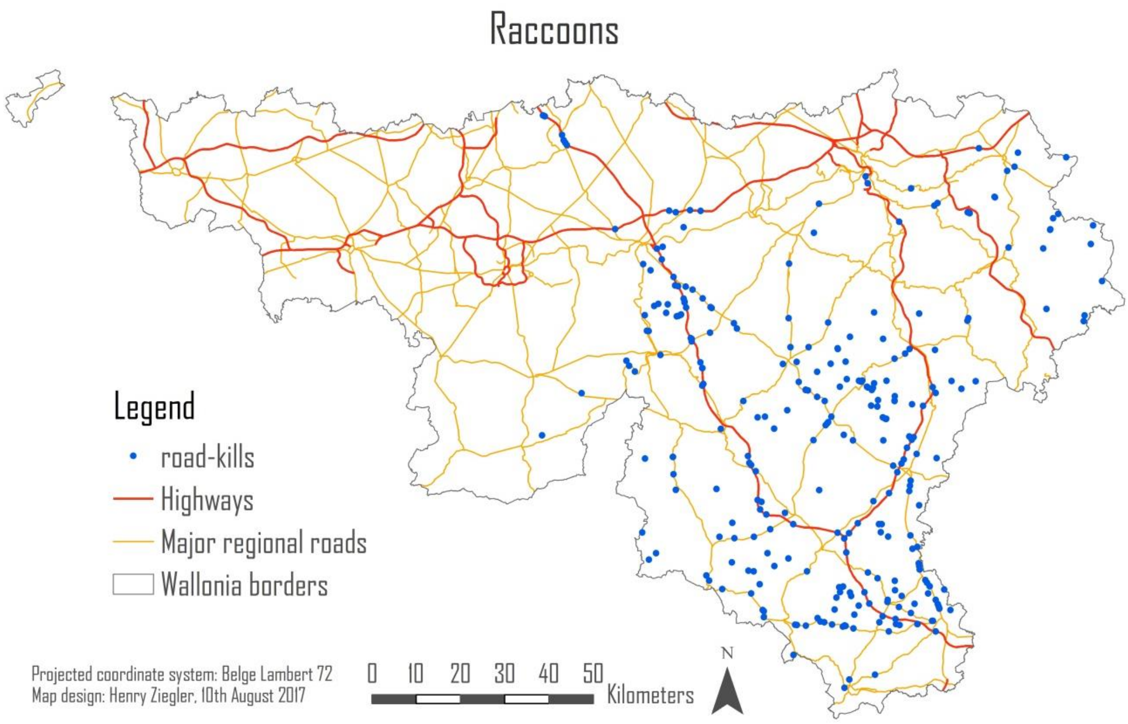


Figure 15 : Map of raccoons' road-kills observations for the period 2006 - 2016

Badgers' observations mainly covered the southern part of Wallonia (South of the Sambre and Meuse rivers). Lots of data were recorded along highways and major roads. On the opposite, polecats' data were more scattered on the whole territory. Collisions occurred with lower densities. Finally, raccoons' carcasses were mainly found in the Ardennes region, at low densities as well.

3.3.2 / Annual numbers of observations

Here are bar-charts of the annual amount of roadkill observations and of observers that record them (Figures 16, 17 and 18)

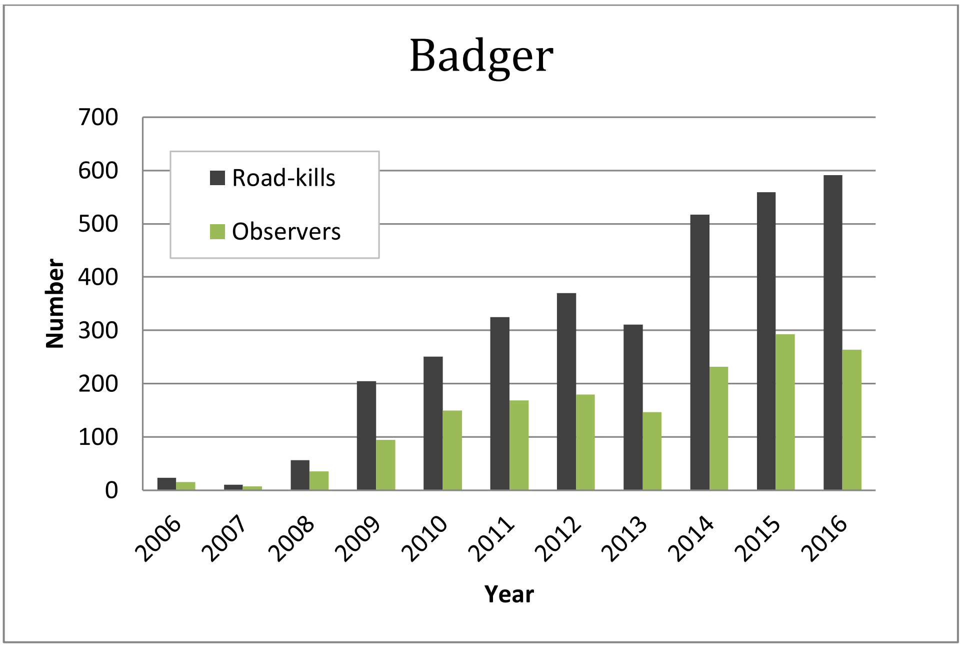


Figure 186 : Annual numbers of badgers' roadkill observations and of observers that record them. (black: roadkills; green: observers)

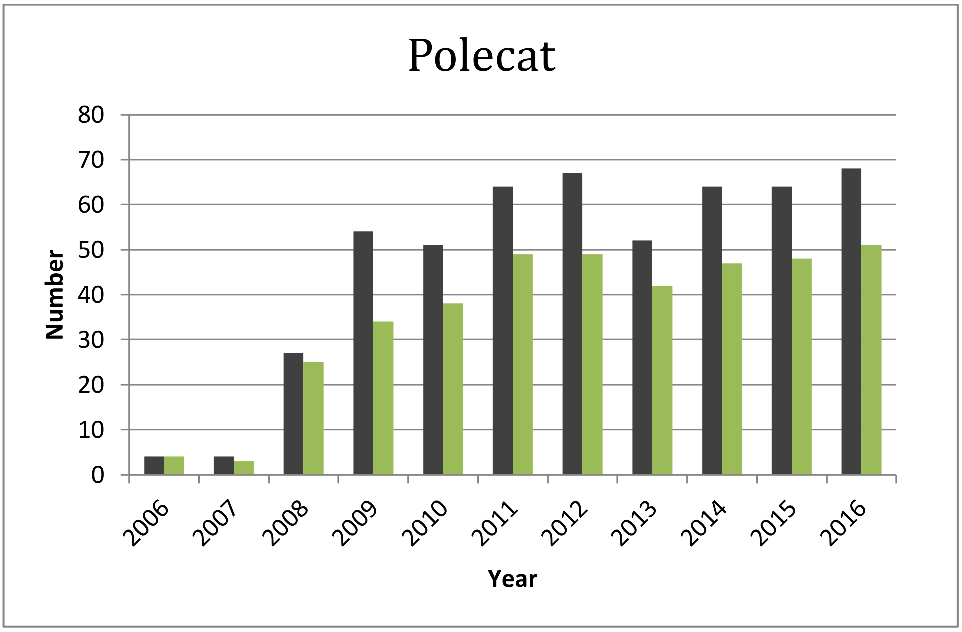


Figure 17 : Annual numbers of polecats' roadkill observations and of observers that record them (black: roadkills; green: observers)

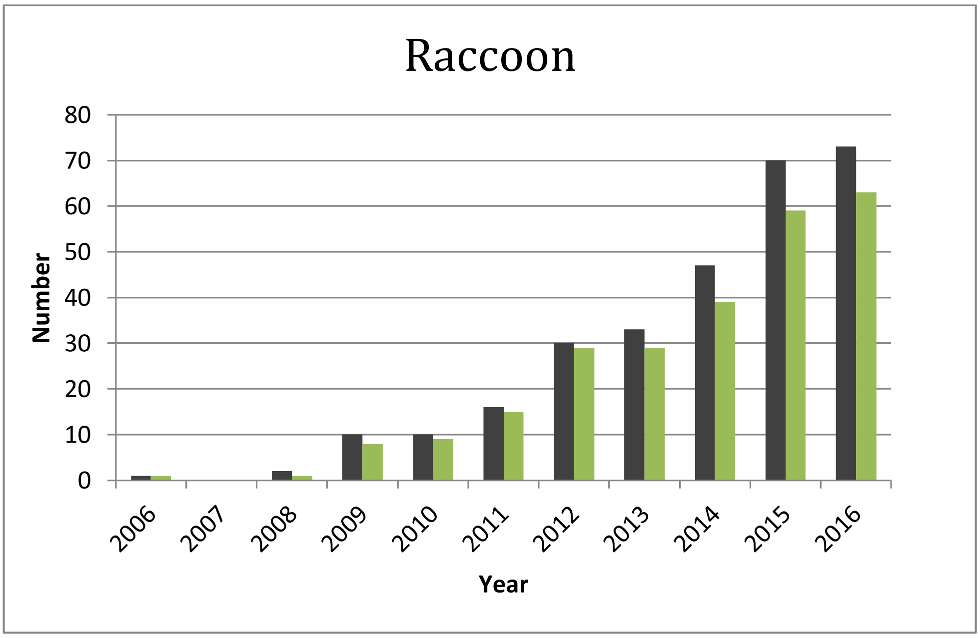


Figure 18 : Annual numbers of raccoons' roadkill observations and of observers that record them (black: roadkills; green: observers)

There were much more badgers annually killed than polecats or raccoons. Numbers of observations and authors increased between 2006 and 2016 for badgers and raccoons (with a striking increase for raccoons). For polecats, these figures remained quite stable since 2009 or 2011. It seemed that observations and authors followed the same trends. To quantify that, correlations have been calculated:

Table 11 : Correlations between the annual numbers of roadkill observations and the number of observers that record them. Those correlations have been calculated for the three study species using two different periods of time.

	Badger	Polecat	Raccoon
2006-2016	0,989	0,988	0,998
2009-2016	0,968	0,896	0,998

Whatever the species or the time period, there were strong correlations between observations and observers.

3.3.3 / Analyses for polecats and raccoons.

Temporal similarities

Here are bar-charts that show the number of traffic victims and the number of other types of data recorded annually (Figures 19 and 20)

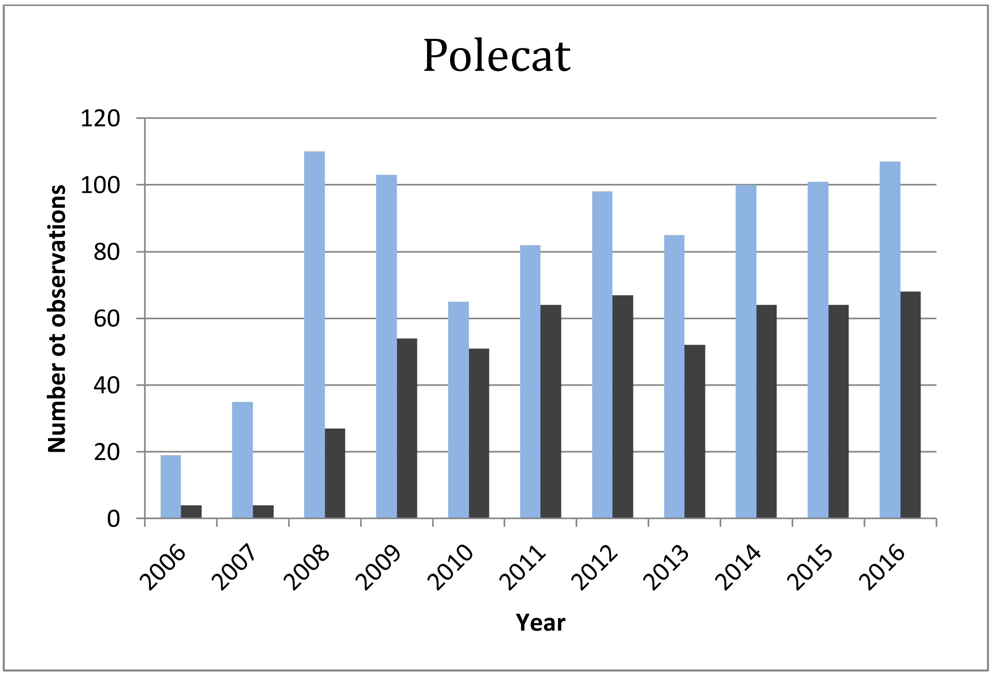


Figure 19 : Polecat's numbers of traffic victims and numbers of other types of data recorded annually. (blue: other data; black: roadkill data)

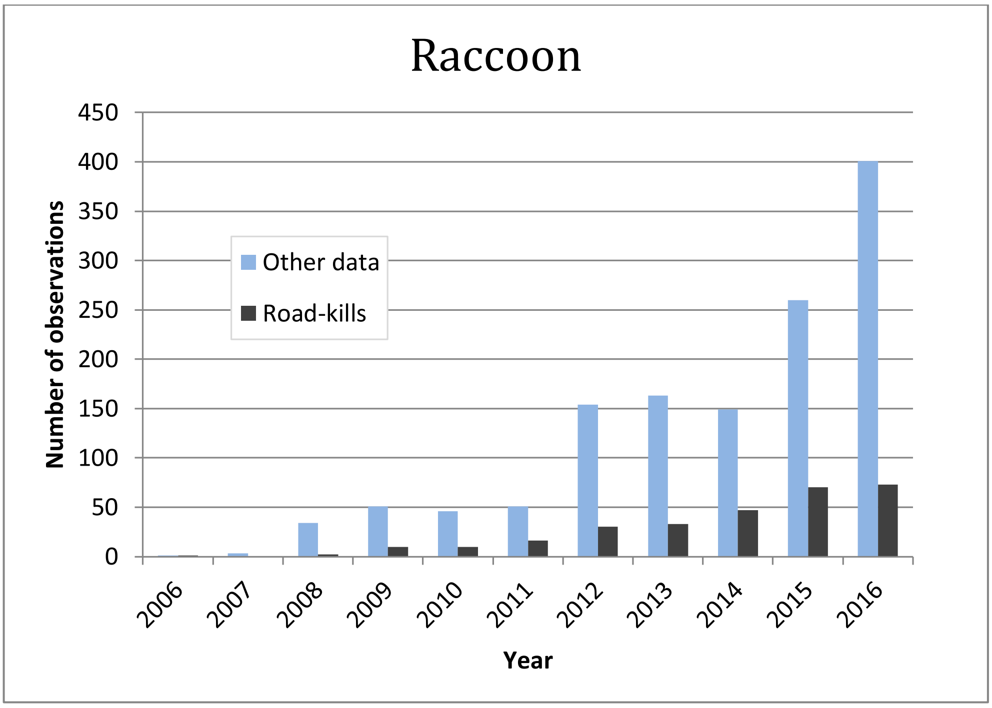


Figure 20 : Raccoons' numbers of traffic victims and numbers of other types of data recorded annually. (blue: other data; black: roadkill data)

In polecat's case, the number of "other data" was slightly higher than the number of collision data. They varied around 80 observations per year. In raccoon's case, "other data" were clearly more numerous and increased regularly. Correlations between the number of "other types" of observations and the number of authors that record them were quite good for polecat, and very high for raccoon (see Table 12 bellow).

Table 12 : Correlations between a) the annual numbers of "other types" observations and the number of observers that record them; b) the annual numbers of roadkill observations and the annual numbers of "other types" observations; c) annual roadkill indexes and annual indices of "other types" observations. These correlations have been calculated for polecats' and raccoons' data using different periods of time and an aggregation of data over 6 years.

Correlation tested	Period	Polecat	Raccoon
N. obs. other types - n. authors other types	2006-2016	0,891	0,989
	2009-2016	0,753	0,986
N. obs. roadkill - n. obs. other types	2006-2016	0,776	0,954
	2009-2016	0,609	0,935
	6-years aggregation	0,939	0,994
Roadkill index - other types index	2006-2016	0,142	0,347
	2009-2016	-0,116	0,347

Correlations between the annual numbers of roadkill observations and the annual numbers of "other types" observations were medium in polecat's case. They were way better when aggregating data over 6 years. Correlations were very high in every case for raccoon.

To take into account the fact that observations and observers were generally well-correlated, indexes have finally been calculated. Their annual variations are displayed in Figure 21 and 22. Correlations between indexes were very poor in every case (Table 12 above).

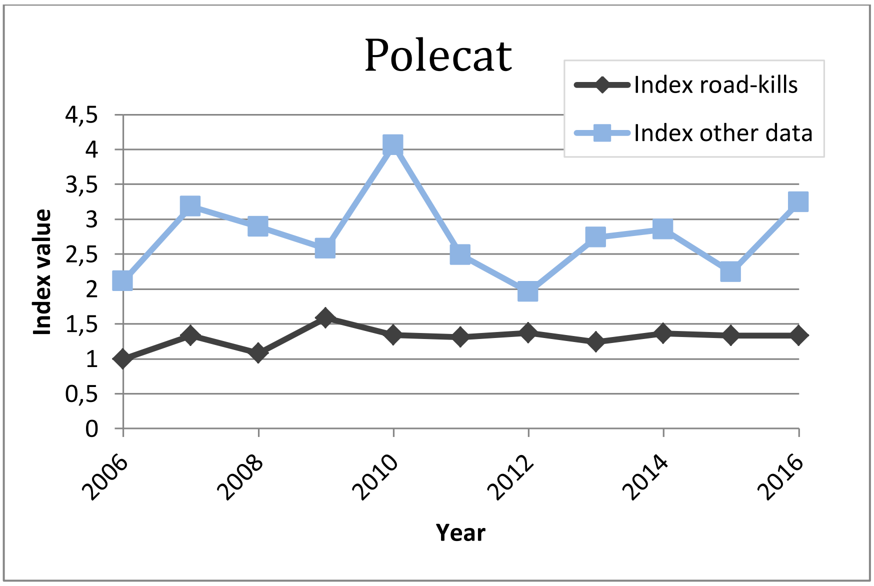


Figure 21 : Annual variations of indexes obtained by dividing the number of observations by the number of observers.

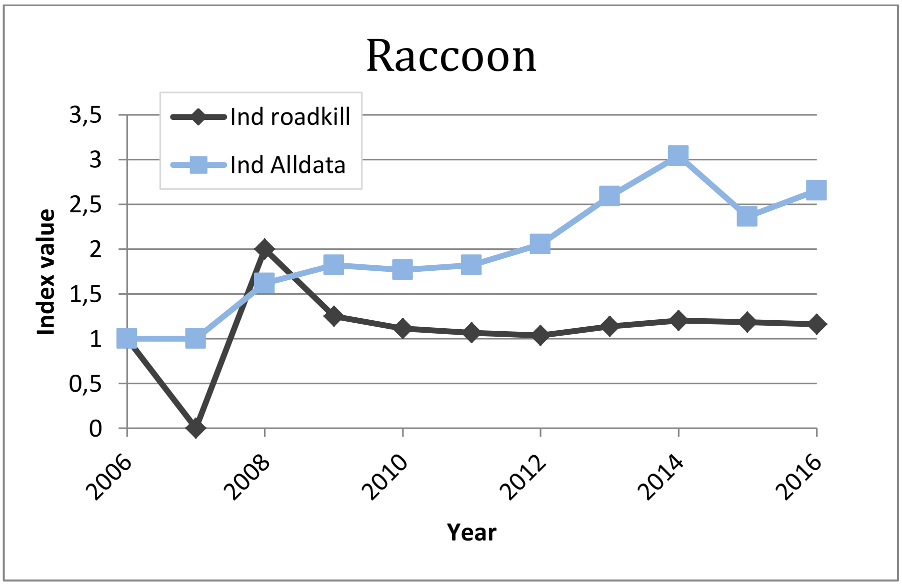


Figure 22 : Annual variations of indexes obtained by dividing the number of observations by the number of observers.

Spatial similarity

Here are the annual numbers of cells occupied by the different categories of data (Table 13):

Table 13 : Annual numbers of 10 x 10 km cells where roadkill or other types of data have been recorded.

	Polecat		Raccoon	
Year	**Roadkill data**	**Other data**	**Roadkill data**	**Other data**
2006	4	15	1	1
2007	3	23	0	3
2008	21	53	2	18
2009	34	49	7	22
2010	41	41	9	24
2011	47	50	13	25
2012	49	60	20	37
2013	38	49	20	47
2014	48	52	31	47
2015	51	55	43	60
2016	46	49	46	73
Total 2006-2016	149	161	79	102

And here are the results of the spatial similarity test (Table 14). To make this table clearer, a scheme has been drawn to illustrate the different values obtained for polecat (Figure 23).

Table 14 : Mean proportions of 10 x 10 km cells that contained a) both roadkill data and other types of data; b) only roadkill observations c) only other types of observations. The mean ratio "common occurrence / number of cells where other types of observations were recorded" is also shown. These mean proportions have been calculated using annual data or data aggregated over 6 years.

	Polecat		Raccoon	
	Annual data	**Data aggregated over 6 years**	**Annual data**	**Data aggregated over 6 years**
Common occurrence	16,91%	52,33%	30,60%	52,66%
Only roadkill observations	37,52%	21,88%	13,92%	13,60%
Only other types of observations	45,57%	25,79%	55,48%	33,74%
Com. Occ. / Total other types	27,00%	66,99%	35,63%	60,92%

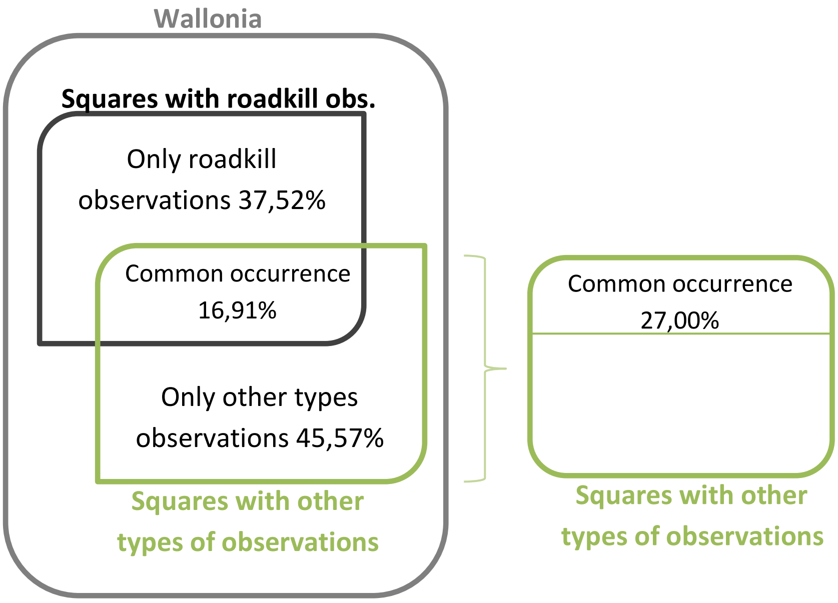


Figure 23 : Scheme of the different types of data occurrence in a 10 x 10 km grid. Values are average proportions calculated with polecats' data over the period 2009-2016. In black: the whole squares containing collision data; in green: the whole squares containing other types of data.

On average, among all the cells where any kind of data has been recorded, only 16,9% and 30,60% (for polecat and raccoon respectively) of these cells contained roadkill observations and other types of data at the same time. On average, only 27,00% and 35,63% of the cells containing "other data" also contained collision data. When summing data over 6 years, those proportions increased to 66,99% and 60,92% respectively.

3.3.4 / Analyses for badger

Among the 3705 badgers' roadkill observations available between 2006 and 2016, 2962 fell within the scope of the model. Data distribution is presented in Figure 24.

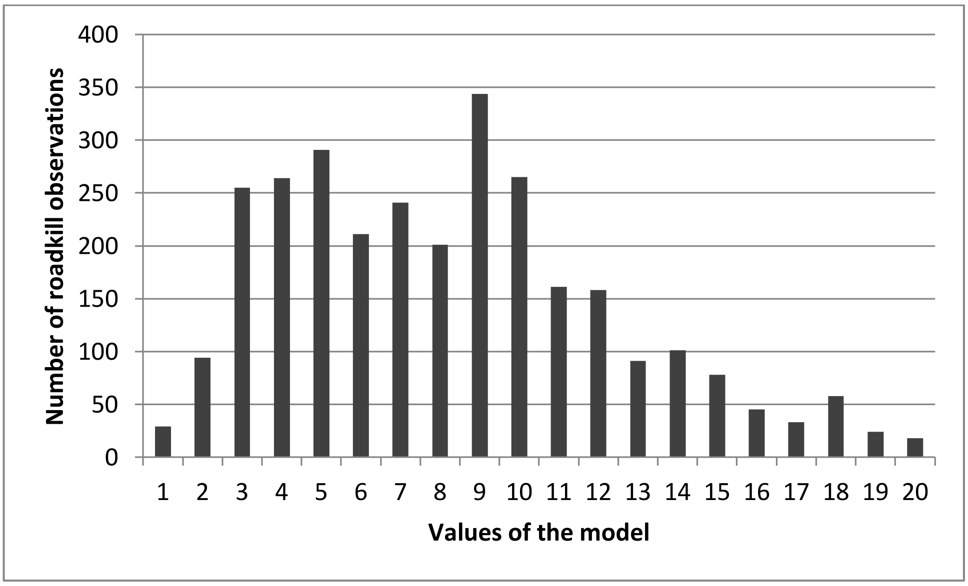


Figure 24 : Number of badger's roadkill observations per value of the model.

This chart for example means that 344 observations had a median model value of nine in their surroundings. The majority (74,11%) of data fell within cells with a model value inferior or equal to ten. 31,50% fell within cells with a model value inferior or equal to five.

3.4 / DISCUSSION

3.4.1 / Maps of roadkill observations

Carcasses' locations corresponded to known species' distributions: badgers in the southern part of Wallonia, polecats on the whole territory and raccoons in the Ardennes.

The fact that lots of badger's data were recorded along highways and major roads is striking. Some roads were covered by road-kills and completely appeared in blue on the map. However, due to unknown sampling efforts, it is hard to determine whether this high number was due to more abundant populations there, to roads characteristics or to a local higher search effort.

In their paper "Traffic mortality and the role of minor roads", van Langevelde et al. (2009) showed that there were more badgers killed on minor roads than on major roads in the Netherlands. However, fences fitted out along highways in the Netherlands are supposed to be impermeable to badgers (which is not the case in Wallonia), which could explain this lower amount road-kills on highways. In Wallonia, it is therefore likely that traffic mortality is high on major and on minor roads at the same time (Vinciane Schockert, pers. discuss.).

3.4.2 / Annual numbers of observations and observers

The number of badgers annually killed was much higher than polecats' and raccoons' numbers. Those figures depend on population status, traffic, behaviour of observers, but also on carcass detectability. For example, the high number of dead badgers encoded could be influenced by the fact that this species is quite easy to detect and recognize when driving. On the opposite, polecats have a small size and are sometimes not easy to recognize. Moreover, they will be degraded faster (Santos et al., 2011). This could explain the lower number of observations.

Number of carcasses' observations increased with years. However, it is likely that this increase in numbers is –at least partly- related to a greater use of databases. Indeed, for example in badger's case, conditions (environment, road infrastructures, populations' status,…) didn't change significantly over past years (event if populations are slightly expending in some areas).
In raccoons' case, the rapid increase in number of observations could be due to a larger use of database and to its invasive expansion at the same time.

Obsevations.be started to run in 2006. Of course, it took time for the public to get to know the existence of this database. Recording of mammals really started only in 2009, sometimes after training sessions (Vinciane Schockert, pers.discuss.). The very good correlation between observations and observers illustrates this influence of databases use. It shows that the annual amount of available data is probably strongly linked to the number of people that pay attention to carcasses along roads.

3.4.3 / Analyses for polecat and raccoon

Temporal similarity

The tiny difference between the number of roadkill data and "other data" for polecats illustrates the importance of collision observations in the total amount of available data. This importance is less clear in the case of raccoon.

When looking at absolute numbers only, it appeared that the correlation between roadkill data and "other data" was usually good (and even very good when aggregating observations over 6 years). On the basis of those correlations only, one would therefore conclude that roadkill data accurately reflect the variations of the rest of available data, which are considered as the reference.

However, as correlations between observations and authors were usually pretty good, it is more relevant to compare roadkill and other data using indexes. In those cases, correlations dropped down to very poor values. As a consequence, it is concluded that temporal similarity was very poor as well. Thus, traffic victims could not indicate temporal variations of populations' status.

Spatial similarity

The number of cells where roadkill data and other data where recorded at the same time was quite low. When using annual data, those situations only accounted for 27,00% and 35,63 of the total number of cells where "other data" have been recorded. This means that traffic victims only reflect the distribution of the rest of the data in some cases.

When summing observations over six years, to figures increased to 66,99% and 60,92% respectively. Of course this is a better result. But still, a better similarity would have been expected when aggregating data over such a long period.

It is therefore concluded that spatial similarity is not sufficient enough. Traffic victims would not be an appropriate indicator of the distribution of other data, which are considered as the reference.

However, there are not useless. Results showed that the number of cells that only contained roadkill data was often high (sometimes the same amount than cells that only contained "other" data). Traffic victims are therefore an important source of information. They are complementary to "other" data and give valuable information about species' distribution.

3.4.4 / Analyses for badgers

Results showed that collisions generally occurred in areas with a good or medium suitability for badgers. 74,11% of the dataset had a median model value of 10 or less in their surroundings. Only few observations occurred in areas with a poor suitability for badgers' burrows.

However, the distribution was not concentrated on best suitability classes only. Collisions didn't always occur in areas with a good potential habitat for badgers. Thus, roadkill data didn't necessarily indicate areas suitable for the mustelid.

Once again, this experiment suffered from the unknown sampling effort. As it is likely that the latter influenced spatial distribution of data, results presented here may not reflect reality.

To improve this experiment, it would be useful to only focus on road sections were search effort is expected to be high (some busy highways sections for example), so that results would be more reliable.

3.4.5 / The issue of sampling effort

Please note that those results depend on the method with which data have been harvested. Since there are often voluntary, it is likely that observations don't perfectly reflect traffic victims and populations' status. Thus, results don't necessarily reflect reality.

This issue mainly depends on the problem of sampling effort. To try to tackle this problem, further work should investigate the relationship between vehicles traffic data and road-kills occurrence. The model designed above should also be tested.

I would also recommend adopting a strategy like the one used in Flanders with the "Dieren Onder de Wielen" project, in order to gather data of better quality in a more organised manner (Vercayie et al., 2015). In this project, some people that regularly drove the same road section agreed to record their daily observations. As a result, sampling effort is known. Those transect data are then used in combination with voluntary observations like those analysed here.

Generally speaking, improving cooperation with the "Dieren onder de Wielen" project would probably be very beneficial. I would strongly recommend exchanging experiences and developing a common project as a common database (observations.be) is used in Flanders an Wallonia at the same time.

3.5 / CONCLUSIONS

In this chapter, roadkill observations have been compared to other data that served as a reference, in order to see whether traffic victims could serve as indicators of populations' status.

Considering the results, it is concluded that the temporal or spatial similarities were too low to consider that this kind of indicator could work. Nevertheless, it has been found out that road-kills represent an important complementary source of information. A greater monitoring of carcasses would therefore be worth.

However, conclusions depend on the way data have been harvested. Changing methodologies would maybe lead to other results and conclusions.

Further works should try to find solutions to tackle the issue of unknown sampling effort, which leads to lots of uncertainties when analysing data. Integrating vehicles' traffic data into analyses and using the model designed above could help solving this problem. Adopting a strategy like the one developed in the "Dieren Onder de Wielen project" would also enable gathering more reliable data.

4 / IMPACT OF TRAFFIC ON POPULATIONS

The third objective of this master thesis was to investigate how roads impact wildlife populations. For that purpose, three operations have been done. There are explained here bellow.

4.1 / METHODS

4.1.1 / Seasonal patterns in traffic mortality

Monthly distributions of collisions' occurrence have been calculated. It seemed relevant, as it is known that the amount of animals killed along roads varies depending on the season (Haigh, 2012). It was therefore interesting to see whether the patterns of our data corresponded to existing literature and fields' observations.

To calculate those monthly distributions, all roadkill observations recorded between 2009 and 2016 have been used in badgers' and polecat's case. For raccoon, only the period 2012-2016 has been selected. Indeed, there were too many months without any observations before 2012. This would have led to results distortion.

Then, for each year, the proportion of individuals killed each month has been calculated. Mean proportions have finally been obtained. Annual percentages have also been displayed as boxplots.

4.1.2 / What proportion of badger's population is killed on roads each year?

To investigate the impacts of road mortality on wildlife populations, it is relevant to try to estimate the portion of the total population that is annually killed on the tar (Dekker et al., 2010). However, we have already said that accurate figures of wild mammals' abundance were tricky to obtain. Generally speaking, there is therefore no precise estimation of the number of individuals that live within Wallonia's territory.

Nevertheless, this kind of figures exists for badgers. Between 2009 and 2014, Walloon badger's population size has been estimated by the Zoogeography Research Unit of the University of Liège. Those estimations were based on burrow surveys and on average numbers of badgers that lived within burrows and clans[17]. Some adjustment factors were also applied to produce final figures.

[17] For more information, see : Schockert V., Lambinet C., Libois R., 2014. *Convention SPW/ULg sur 15 espèces de mammifères protégés ou concernés par la Convention de Berne et sur 2 espèces envahissantes. Rapport d'activités 2014.* Liège : Unité de recherche zoogéographique, Université de Liège. (internal report, unpublished)

Populations' estimations have then been summarized and compared to the amount of roadkill observations recorded each year on roads. An average proportion of badgers that were killed annually has finally been calculated.

4.1.3 / Locating hotspots of collisions

Many studies confirmed that collisions are spatially clustered and that some "hotspots" of animal mortality can be identified (Clevenger et al., 2003; Morelle et al., 2013; Barthelmess, 2014, etc.). Locating potential or actual hotspots therefore enables identifying issues and designing mitigation measures to reduce roads' impacts.

Here, it has been decided not to only locate those hotspots (which would be a too restrictive approach) but rather to classify zones using a gradient that ranges from "least concern" to "high mortality area". This would allow locating, from one hand, high or medium priority zones and from the other hand areas where collisions are less important.

To do so, UTM grids with cells of 1 x 1 km and 5 x 5 km have been used. These cells sizes enabled studying road-kills densities at different scales. The goal was also to design maps that could be easy-to-use and that could serve as tools to tackle the issue of wildlife traffic mortality.

Then, the mean number of collision data recorded each year has been calculated for each cell using "Spatial Join" tool in ArcGIS. Observations encoded between 2009 and 2016 have been used only. Indeed, the little amount of data recorded before 2009 would have influenced means too much.

The question then was to determine how to classify grid cells depending on their mean number of road-killed animals. This was an important point as different classifications could lead to very different maps. Astonishingly, almost none of the mentioned papers defined what they called "hotspots". They simply displayed road networks and densities of mammal mortality. As there was no reference to be based on, it has been decided to split the distribution of the mean number of road-kills per square into five classes using five equal intervals (Figure 25). The upper class corresponds to high mortality areas, the lower one to zones with only few animals road-killed.

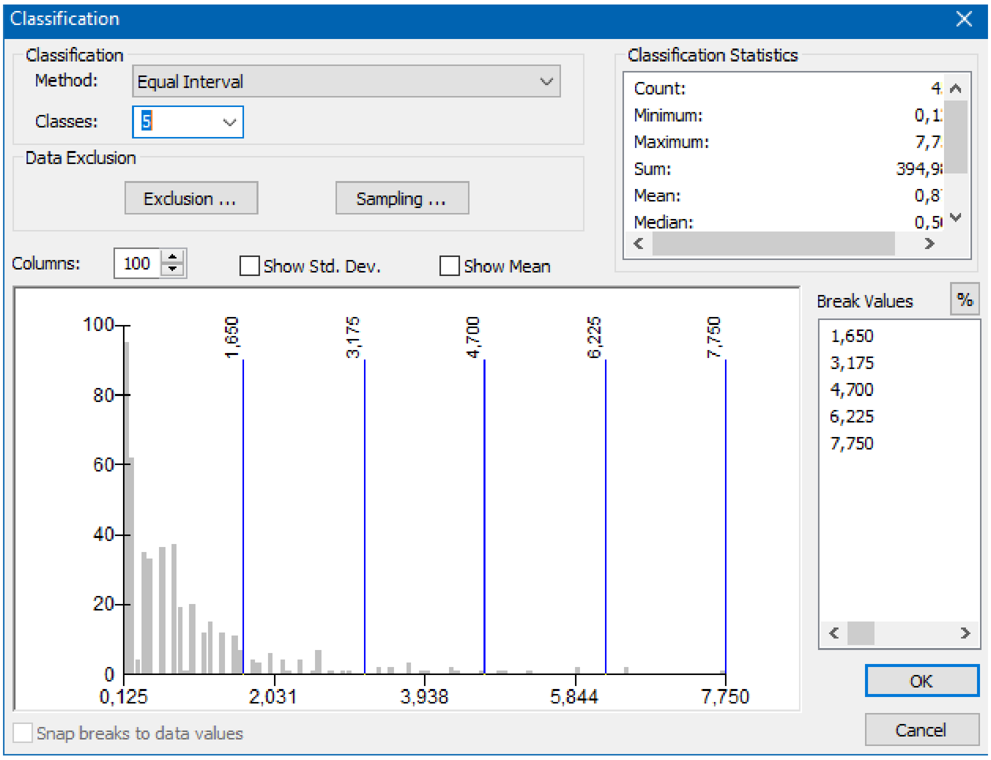


Figure 25 : Classification used to display mean numbers of traffic victims per 5 x 5 km cell in ArcGIS.

As locating collisions hotspots is rather a conservation approach intended to protect species from roads' impacts, hotspots have only been identified for badgers and polecats. Indeed, conservation issues are associated to these species, on the contrary to raccoon that is an invasive species.

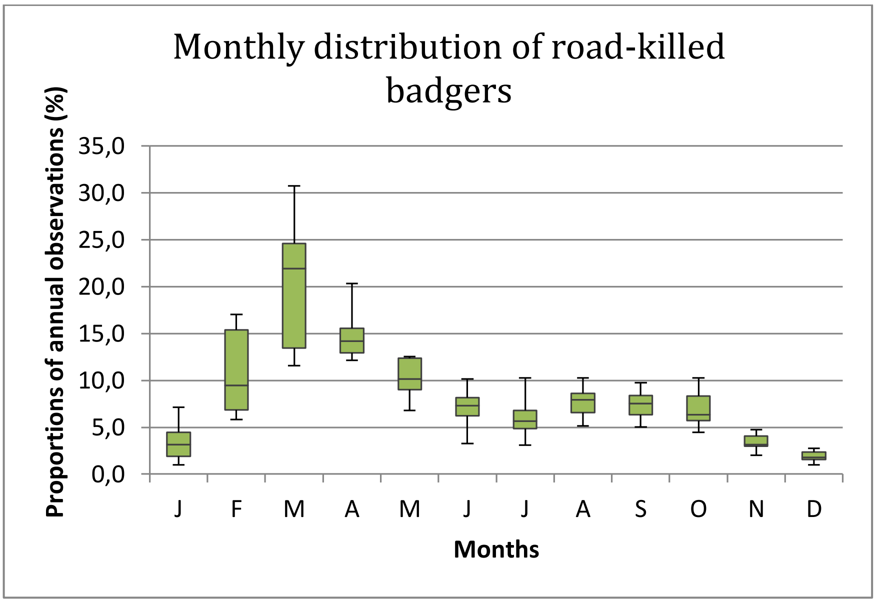
4.2 / RESULTS

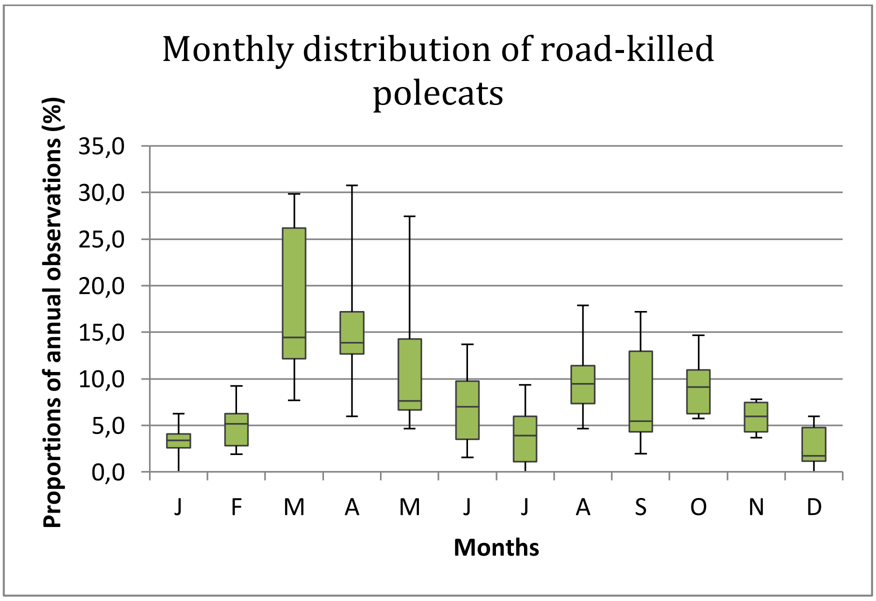
4.2.1 / Seasonal patterns in traffic mortality

Calculating monthly distributions of road-killed animals gave the following results (Table 15 and Figure 26)

Table 15 : Mean proportion (%) of each month in the total amount of road-killed badgers, polecats and raccoons.

Month	J	F	M	A	M	J	J	A	S	O	N	D
Badger	3,4	10,8	20,4	14,7	10,2	7,2	5,9	7,7	7,5	6,9	3,4	1,9
Polecat	3,3	4,9	18,0	15,8	11,2	7,0	3,9	9,9	8,4	9,2	5,9	2,7
Raccoon	1,9	6,2	3,6	5,4	5,8	4,7	6,0	13,4	13,3	24,2	9,8	5,7





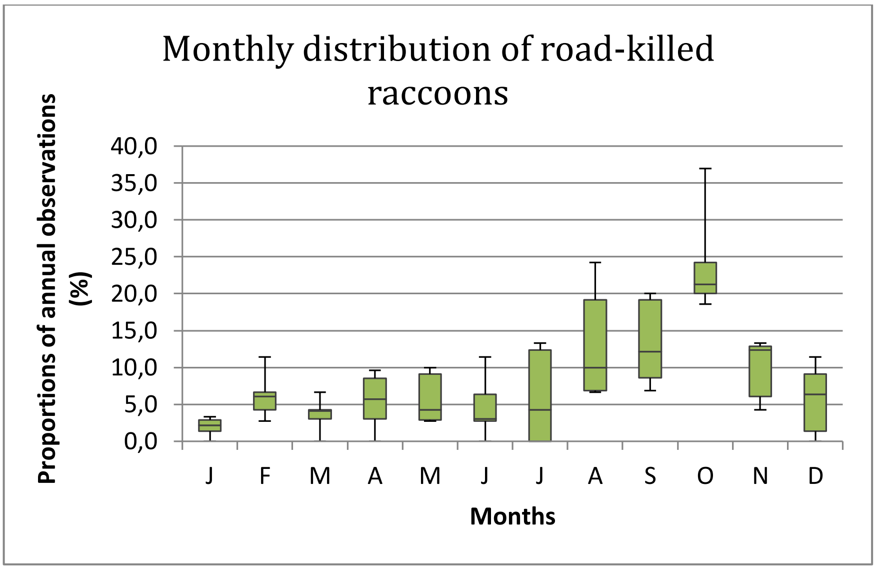


Figure 26 : Boxplots of the monthly distribution of road-killed badgers, polecats and raccoons. Quantities are expressed in proportions of annual observations (%).

Badger's distribution showed one main peak. Indeed, collisions mainly occurred in spring, especially in March. Polecat's distribution was more scattered, with a higher variability within months. There were two peaks: one in spring and the other in late summer. Finally, raccoon's distribution showed a clear increase in road-kills in late summer/autumn.

4.2.2 / What proportion of badger's population is killed on roads each year?

Here are the estimated proportions calculated (Table 16):

Table 9 : Comparison between estimations of badgers' populations' levels and the annual number or roadkills.

	Estimated number of individuals		Roadkill data	
Year	**Min.**	**Max.**	**Number**	**Min. - Max. proportions (%)**
2009	2988	5713	205	3,59 - 6,86
2010	3260	5579	251	4,5 - 7,7
2011	3211	5834	325	5,57 - 10,12
2012	3532	6018	370	6,15 - 10,48
2013	3379	5923	311	5,25 - 9,2
2014	3777	6964	517	7,42 - 13,69
Mean	3357,8	6005,2	329,8	5,41 - 9,67

So, the average annual proportion of road-killed badgers reaches at least 5,41% of the global population. However, this figure is probably underestimated. It is very likely that the true number is much greater (see discussion paragraph).

4.2.3 / Locating hotspots of collisions

Mapping average number of carcasses recorded each year per cell enabled producing the following maps (Figures 27, 28, 29, 30).

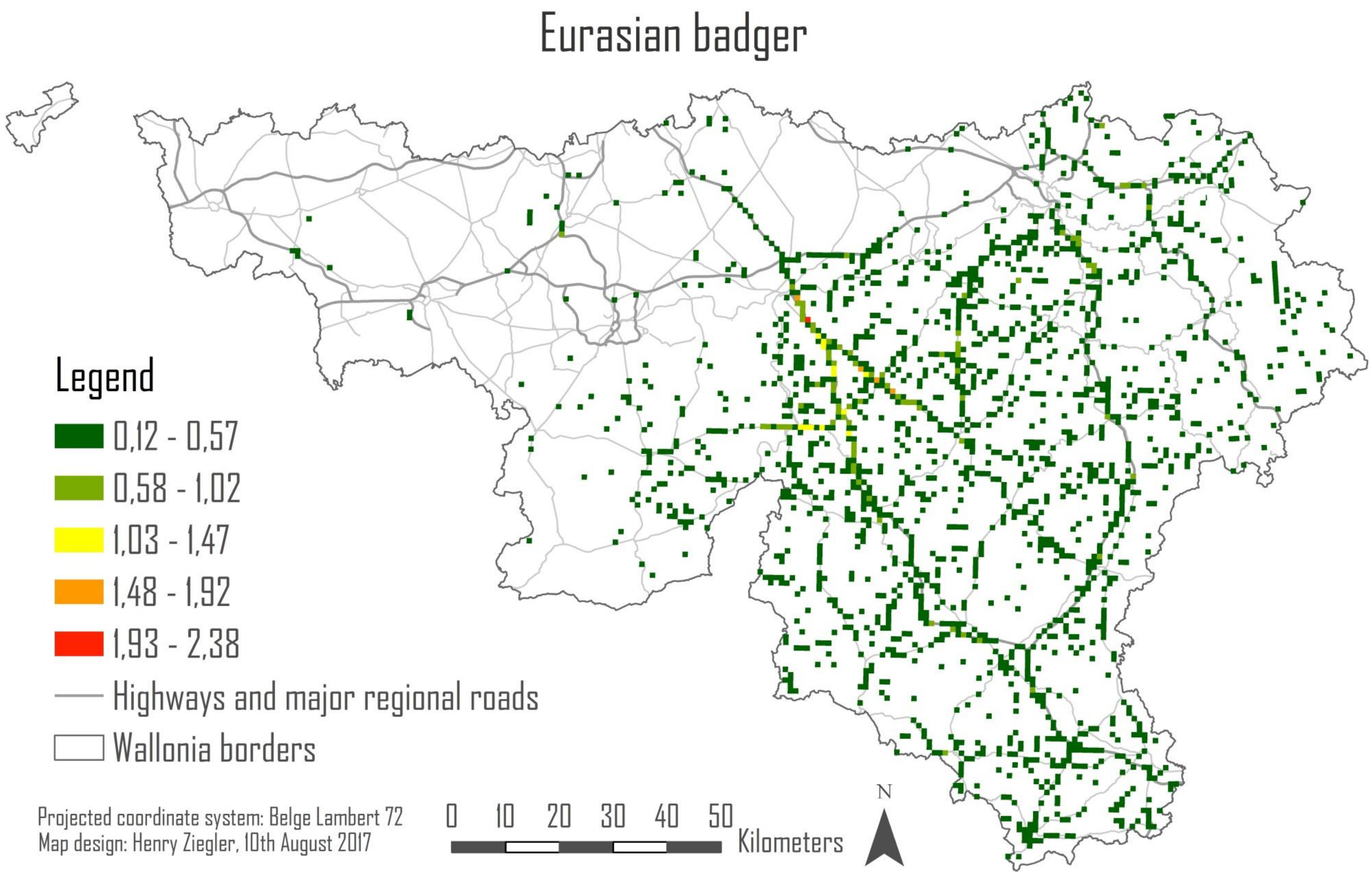


Figure 27 : Map of the average number of badgers annually road-killed within 1 x 1 km grid cells.

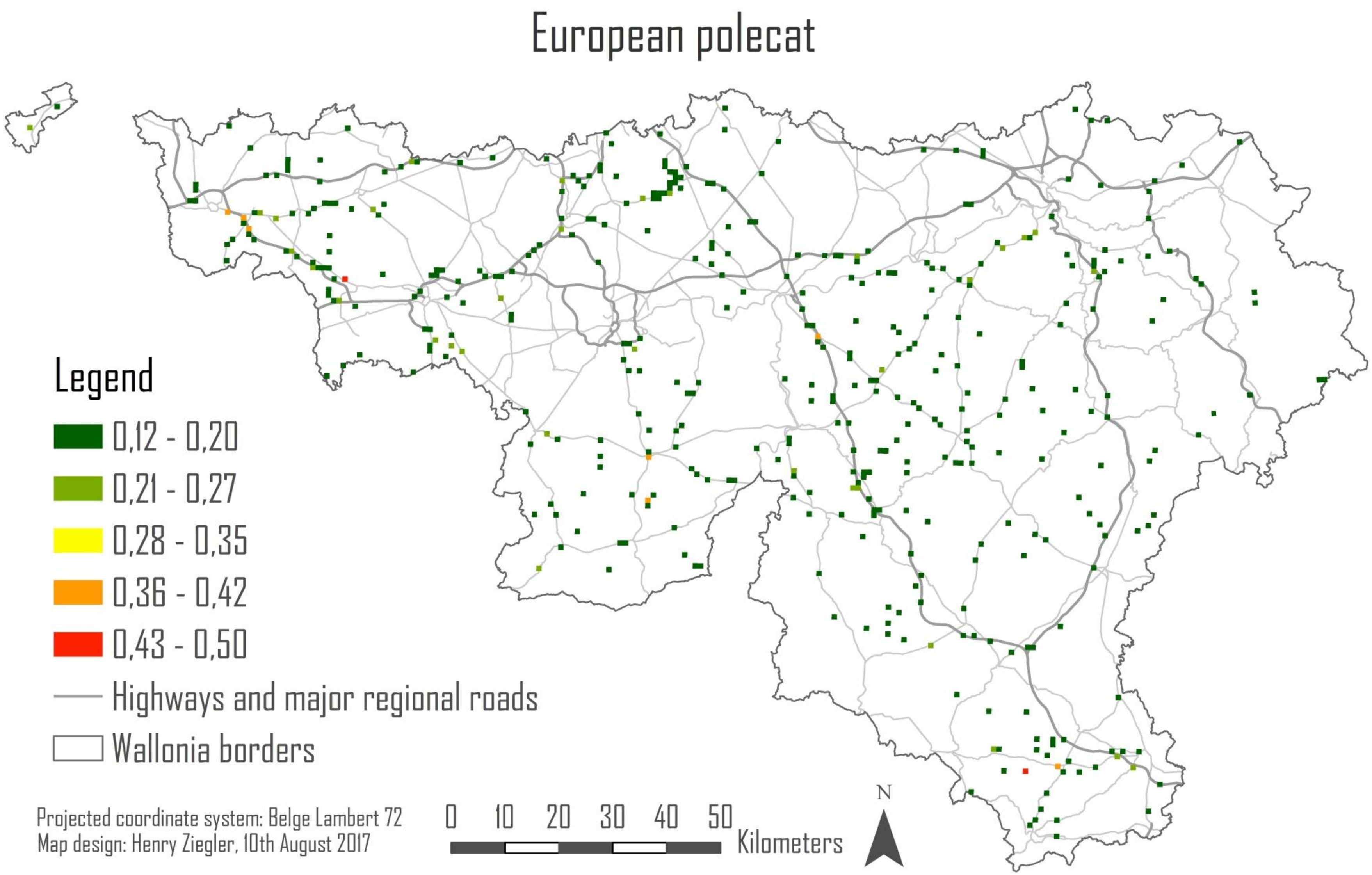


Figure 28 : Map of the average number of polecats annually road-killed within 1 x 1 km grid cells.

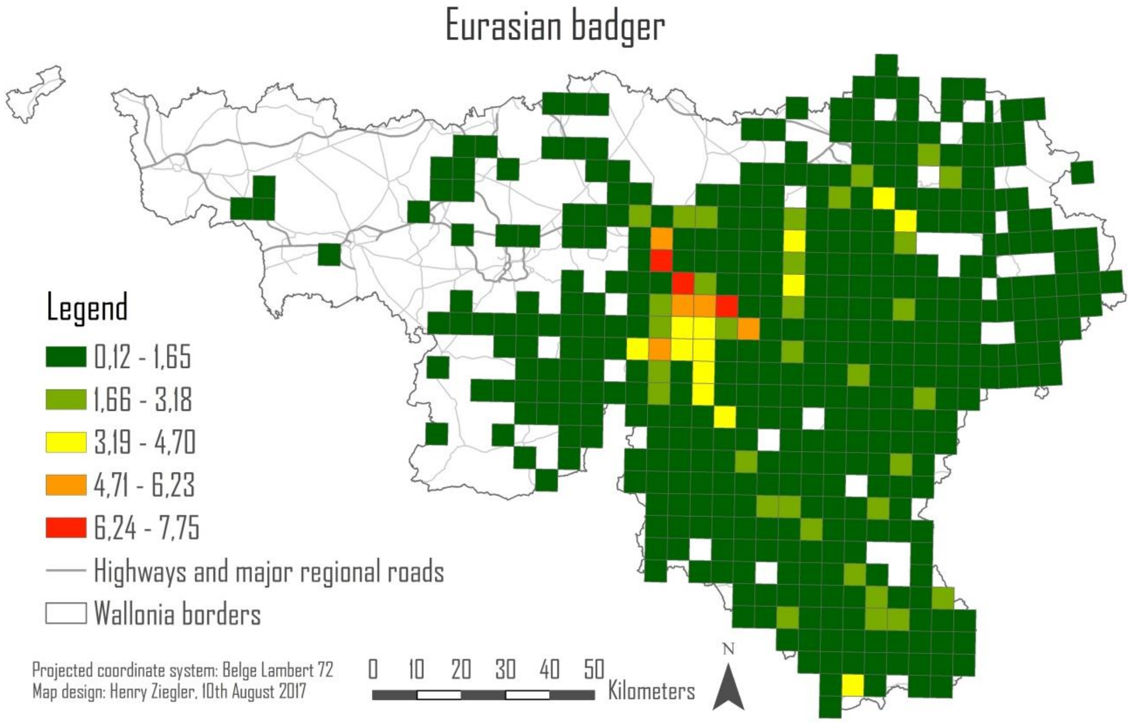


Figure 29 : Map of the average number of badgers annually road-killed within 5 x 5 km grid cells.

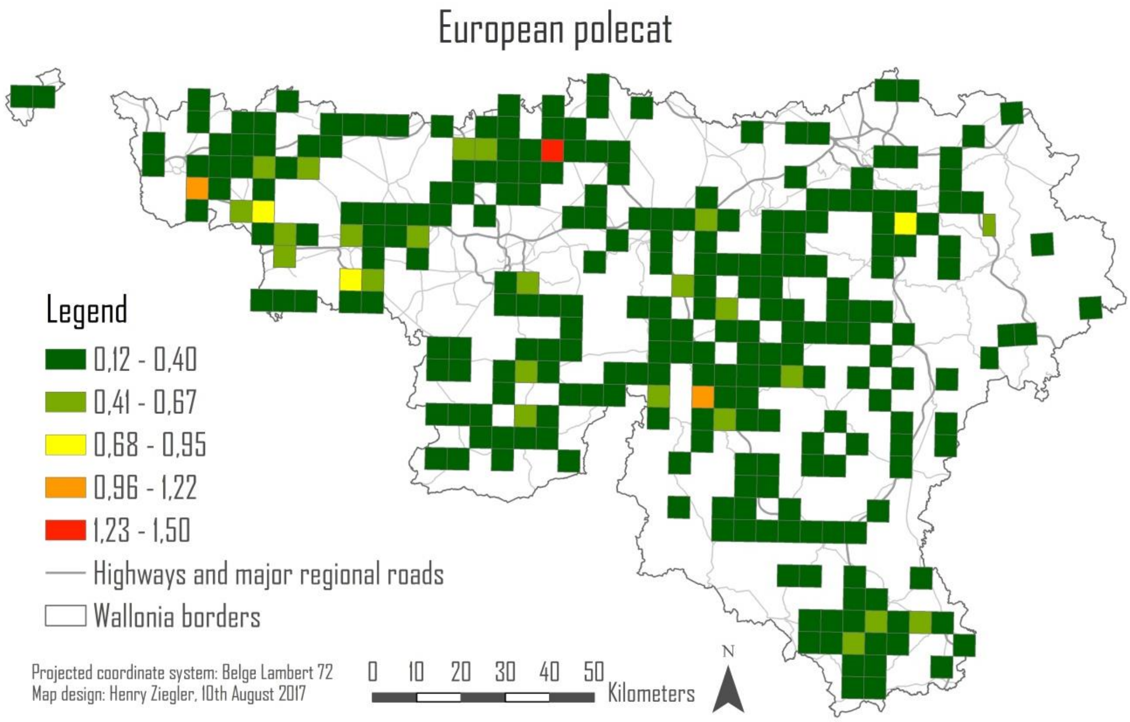


Figure 30 : Map of the average number of polecats annually road-killed within 5 x 5 km grid cells.

It can be observed that spatial distributions of hotspots are not the same for the two species. Badger's hotspots were mainly located in the surroundings of the cities of Namur, Ciney Marche-en-Famenne and Dinant. On the opposite, polecat's hotspots were rather located in the northern part of Wallonia. Nevertheless, other hotspots could also exist (see discussion section).

4.3 / DISCUSSION

4.3.1 / Monthly patterns

Here, the issue of unknown sampling effort was less important. Indeed, it is likely that search effort has almost been the same all years long (Diemer Vercayie, pers. discuss.). It is therefore expected that those results are representative of the real situation (for once…)

Badger's distribution matched existing literature and field observations. For example, Dekker et al. (2010), Clarke et al. (1998), Haigh (2012), observed the same patterns. They even sometimes found a second peak in late summer. Those studies reported that the increase in badgers' activity and mobility due to dispersal of young adults would probably be the cause of this peak in spring.

No information has been found in literature regarding seasonal patterns of polecats and raccoons.

4.3.2 / Proportions of badgers killed annually

It has been found that at least 5,41% of badger's population is killed each year on road. However, this percentage is a very minimum and is probably largely underestimated:

- Firstly, because this percentage is a mean, calculated by taking into account early years of databases, when less people encoded observations. It is probable that the years 2014, 2015 and 2016 reflect better the actual number of flattened badgers. When using data of 2014 only, the proportion of traffic victims then varies between 7,42% and 13,69%.

- Secondly, because populations' estimations use adjustment factors that increase the number of living badgers. If those adjustment factors are not taken into account, then the proportion of traffic victims in 2014 varies between 15,40% and 28,40%

- Finally, because it is likely that many road-killed badgers are not recorded.

In the Netherlands, Dekker et al. (2010) estimated that "between 12% and 18% of the Dutch badger population was killed by traffic each year". They also mentioned those percentages were probably underestimated. Those figures correspond to those found here.

As a consequence, an important part of badger's population is killed on roads each year. Taking action to reduce the amount of victims of this protected species would be definitely worth.

4.3.3 / Locating hotspots

To define classes of collision intensities, the ideal solution would have been to know the thresholds above which wildlife populations are at risk. A hotspot would therefore have been defined as an area where the number of casualties was greater than this threshold. However, those thresholds are unknown, and they depend on local species' abundance. The solution of equal intervals has finally been adopted.

The number and significance of hotspots depends on criteria that define them. Particularly, colour classes used to design maps are very sensitive to these criteria. For further researches, I would therefore recommend to clearly define what a hotspot is and to stick to the same colour classes in order to avoid confusions.

Using a 1x1km grid allowed locating precisely where collisions hotspots were, whereas a 5x5km grid was useful to get an overview of the situation. Polecat's case reveals that when collision data are homogeneously scattered over the whole territory, then it is less worth using the 1x1km grid. Indeed, most of the squares only contain one roadkill observation. The 5x5km grid would be more useful in this case.

Many of the studies I have read about roadkill monitoring used complicated spatial network analysis to produce hotspots maps. This sometimes requires GIS extensions and skilled researchers to perform such analysis. I here argue that using a grid is a much more easy solution that could still deliver useful results.

Because squares have been used to summarize data and to find out hotspots, it is not possible to know exactly what roads segments are problematic by simply looking at the grid map (a square encompasses several segments of roads of different types). Nevertheless, if somebody wants to analyse more precisely where areas of collisions located are, he/she would just need to zoom in the red squares, display roadkill points, and see where they appear exactly. Thus, using a grid offers a simple and pragmatic manner to produce hotspot maps and to get an overview of territory's situation.

Since available roadkill observations have been harvested using voluntary sources, data can only be interpreted as "presence" data. In this case, it means that one could not pretend there is no hotspot elsewhere. Actual hotspots significances could also be different from those calculated in our maps. Nevertheless, these maps could still be used as a tool to identify areas where collision densities are very high and where mitigation measures could be fitted out.

4.3.4 / Final comments

In his speech (13th February 2013), Walloon Minister DiAntonio said that investments in passages for wildlife are questionable as collisions with fauna don't occur so often[18]. The Minister was apparently misinformed when saying that. This master thesis clearly showed that a large number of animals are killed every year on Walloon roads.

As a reminder, a collision generally leads to death for the animal, but also to damages for vehicles. 500 to 600 badgers at least are killed every year nowadays in Wallonia, which potentially corresponds to the same amount of damaged vehicles. Taking action is therefore not only in the interest of fauna but also of drivers. There are already several very good reports and handbooks about designing mitigation measures, namely those published by the COST 341 ACTION of the Infra Eco Network Europe[19]. They could serve as guidelines for future projects.

Last but not least, a new highway segment is currently being built around the city of Couvin, south-west Wallonia. Fauna over and underpasses are implemented there in the meantime. This is a great opportunity to test the efficiency of this kind of passages and to study road permeability to wildlife. I would therefore recommend conducting studies there.

4.4 / CONCLUSIONS

In this chapter, several statistics related to the way roads impact wildlife have been produced.

Monthly distributions enabled getting an idea of the seasonal patterns of collisions. In particular, Badger's distribution showed an important peak in spring, which matches existing literature's results and field observations.

The proportion of badgers annually killed depended on estimations. A very minimum value of 5,41% has been obtained. However, it is very likely that actual percentages are higher (for example 7,42% to 13,69% or 15,40% to 28,40%, or even higher, depending on calculation method).

Finally, hotspots maps have been produced. Even if there are only partially reliable because of the problem of sampling effort, they could be used as a tool to identify dense areas of collisions.

It is recommended to take action to reduce the number of collisions as it impacts many mammals but also many drivers in the meantime.

---

[18] https://www.parlement-wallonie.be/pwpages?p=interp-questions-voir&type=all&id_doc=44921 (14/08/2017)

[19] All documents available at : http://www.iene.info/cost-341-action/ (14/08/2017)

5 / OVERALL CONCLUSION

This master thesis was a first attempt to use voluntary roadkill data and to make them valuable. The three research questions initially asked have been answered (at least partially):

- A proper methodology has been designed to sort out and select relevant data. This implied dealing with “specimen” observations, selecting “carcasses” data on the basis of their distance to roads, and removing double-counts.
- In order to analyse those data, sources of biases have been identified and a model that conceptualizes the issue has been created. The similarity between roadkill data and a reference has been tested in order to see whether traffic victims could serve as populations’ indicators. Results showed poor to medium similarity.
- Several statistics have been calculated and hotspots maps have been designed to evaluate to what extent roads impacted wildlife populations. In particular, it has been showed that an important proportion of badger’s population is road-killed every year.

Using this kind of data implied dealing with lots of uncertainties. The most problematic one was probably the absence of coordinated sampling program. As a consequence, search efforts remained unknown. Results presented above are therefore unfortunately only partially reliable.

To take this kind of analysis to the next step, further work would be needed to tackle the issue of sampling effort. For this purpose, the model created above could be used. This would help improving results’ reliability.

It is also recommended to adopt a monitoring program like the one used in Flanders. Recording road-kills on constant transects would enable harvesting data of better quality.

Finally, observations encoding should be improved, in particular by adding a new category in OFFH form and by making observers more aware of the challenges related to dealing with those data. The ideal solution would also be to create a single database to record every biological observation. All those improvements would increase analyses efficiency.

REFERENCES

- Baker P.J. et al., 2004. Is it possible to monitor mammal population changes from counts of road traffic casualties? An analysis using Bristol's red foxes *Vulpes vulpes* as an example. *Mammal Review*, 34, 115–130.

- Barrientos R., Bolonio L., 2009. The presence of rabbits adjacent to roads increases polecat road mortality. *Biodiversity and Conservation*, 18, 405–418.

- Barthelmess E.L., 2014. Spatial distribution of road-kills and factors influencing road mortality for mammals in Northern New York State. *Biodiversity and Conservation,* **23** (10), 2491–2514.

- Clarke G.P., White P.C.L., Harris S., 1998. Effects of roads on badger *Meles meles* populations in south-west England. *Biological Conservation*, 86, 117-124.

- Clevenger A.P., Chruszczc B., Gunsonc K.E., 2003. Spatial patterns and factors influencing small vertebrate fauna road-kill aggregations. *Biological Conservation*, 109, 15–26.

- Coffin A.W., 2007. From roadkill to road ecology: a review of the ecological effects of roads. *Journal of Transport Geography*, 15, 396–406.

- Council of the European Communities, 1992. *Council directive 92 / 43 / EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora*. Official Journal of the European Communities, 22 July 1992, No L 206 / 7–50.

- Dekker J.J.A., Bekker H., 2010. Badger (*Meles meles*) road mortality in the Netherlands: the characteristics of victims and the effects of mitigation measures. *Lutra*, **53** (2), 81-92.

- Devictor V., Whittaker R.J., Beltrame C., 2010. Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions*, 16, 354–362.

- European Parliament and the Council of the European Union, 2014. *Regulation (EU) No 1143/2014 of the European Parliament and of the Council of 22 October 2014 on the prevention and management of the introduction and spread of invasive alien species*. Official Journal of the European Communities, 4 November 2014, L 317/35–55.

- Forman R.T.T., Alexander L.E., 1998. Roads and their major ecological effects. *Annual Review of Ecology and Systematics*, 29, 207–231.

- Frantz A.C. et al. 2012. Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (Cervus elaphus), but not wild boars (Sus scrofa). *Molecular Ecology*, 21, 3445–3457.

- Gehrt S.D., 2002. Evaluation of spotlight and road-kill surveys as indicators of local raccoon abundance. *Wildlife Society Bulletin*, 30, 449-456.

- George L., Macpherson J.L., Balmforth Z. & Bright P.W., 2011. Using the dead to monitor the living: can road kill counts detect trends in mammal abundance? *Applied Ecology and Environmental Research*, **9** (1), 27-41.

- Glista D.J., DeVault T.L., DeWoody J.A., 2009. A review of mitigation measures for reducing wildlife mortality on roadways. *Landscape and Urban Planning*, 91, 1–7.

- Grilo C., Bissonette J.A., Santos-Reis M., 2009. Spatial–temporal patterns in Mediterranean carnivore road casualties: consequences for mitigation. *Biological Conservation*, 142, 301-313.

- Gunson K.E., Mountrakis G., Quackenbush L.J., 2011. Spatial wildlife-vehicle collision models: a review of current work and its application to transportation mitigation projects. *Journal of Environmental Management*, 92, 1074-1082.

- Haigh A., 2012. Annual patterns of mammalian mortality on Irish roads. *Hystrix, the Italian Journal of Mammalogy*, **23** (2), 58–66.

- Holsbeek L., Rodts J. & Muyldermans S., 1999. Hedgehog and other animal traffic victims in Belgium: results of a countrywide survey. *Lutra*, 42, 111–119.

- Jaeger J.A.G., Fahrig L., 2004. Effects of road fencing on population persistence. *Conservation Biology*, **18** (6), 1651 – 1657.

- Jaeger J.A.G. et al., 2005. Predicting when animal populations are at risk from roads: an interactive model of road avoidance behavior. *Ecological Modelling*, 185, 329–348.

- Kristiansen L.V., Sunde P., Nachman G. & Madsen A.B., 2007. Mortality and reproductive patterns of wild European polecats *Mustela putorius* in Denmark. *Acta Theriologica*, **52** (4), 371-378.

- Lehaire F., Morelle K. & Lejeune P., 2013. Collisions entre véhicules et animaux en liberté: état des lieux à partir d'une enquête au sein de la police. *Forêt Wallonne*, 122, 13–21.

- Morelle K., Lehaire F., Lejeune P., 2013. Spatio-temporal patterns of wildlife-vehicle collisions in a region with a high-density road network. *Nature Conservation*, 5, 53–73.

- Olson D.D. et al., 2014. Monitoring wildlife-vehicle collisions in the information age: how smartphones can improve data collection. *PLoS ONE*, **9** (6), e98613.

- Rolley R.E., Lehman L.E., 1992. Relationships among raccoon road-kill surveys, harvests, and traffic. *Wildlife Society Bulletin*, 20, 313-318.

- Rondinini C., Doncaster C.P., 2002. Roads as barriers to movement for hedgehogs. *Functional Ecology*, 16, 504–509.

- Rytwinski T. et al., 2016. How effective is road mitigation at reducing road-kill? A meta-analysis. *PLoS ONE*, **11** (11), e0166941.

- Santos S.M., Carvalho F., Mira A., 2011. How long do the dead survive on the road? Carcass persistence probability and implications for road-kill monitoring surveys. *PLoS ONE*, **6** (9), e25383.

- Seiler A., 2004. Trends and spatial patterns in ungulate-vehicle collisions in Sweden. *Wildlife Biology*, 10, 301-313.

- SPW, 2014. *Les Indicateurs Clés de l'Environnement Wallon 2014 (ICEW 2014)*. Direction de l'Etat Environnemental SPW - DGO3 - DEMNA – DEE.

- Van Langevelde F., van Dooremalen C., Jaarsma C.F., 2009. Traffic mortality and the role of minor roads. *Journal of Environmental Management*, 90, 660-667.

- Vercayie D. et al., 2012. *Monitoring van verkeersslachtoffers langs Vlaamse wegen: "Dieren onder de wielen"*. Rapport van de Vlaamse overheid – Departement Leefmilieu, Natuur en Energie, Natuurpunt Studie vzw en Vogelbescherming Vlaanderen vzw. België.

- Vercayie D., Herremans M., 2015. Citizen science and smartphones take roadkill monitoring to the next level. *Nature Conservation*, 11, 29–40.

- Vercayie D., Lambrechts J., 2017. *Inventarisatie en evaluatie van de impact van het verkeer op wilde dieren in Vlaanderen – "Dieren onder de wielen 2.0"*. Eindrapport. Mechelen: Rapport Natuurpunt Studie. (in press)

APPENDIX 1 : PASSAGES FOR WILDLIFE IN WALLONIA

This list has been established by merging information available at road administration and in the master thesis of Baghli M. (1996) cited above. Contact at road administration (DGO1): Luc Grisard (luc.grisard@spw.wallonie.be), +32 (0)4 231 63 81.

Structure	Type	Official name	Municipality	Start date	Road	X coordinates[1]	Y coordinates[1]	Latitude[2]	Longitude[2]
Overpass	Mixed	Pont n°s43 "passage a gibier" a les tailles	HOUFFALIZE	1989	A26	250670	100070	50,203128	5,778970
Overpass	Mixed	Pont n°s39 "passage a gibier" a la Baraque de Fraiture	LIERNEUX	1982	A26	246722	106515	50,261724	5,725328
Overpass	Specific	Passerelle a gibier à Rulles	HABAY	1988	A4	236946	48205	49,739076	5,574948
Overpass	Specific	Passage a gibier à Grunhaut	LIMBOURG	2004?	A3	260612	149969	50,650389	5,931562
Underpass	Mixed	Unknown	Libin	1985	A4	212547	68729	49,927232	5,238560
Underpass	Mixed	Unknown	LIBRAMONT-CHEVIGNY	1976	N89	221272	72406	49,959304	5,360741
Underpass	Mixed	Buse pour gibier du champ de Harre	MANHAY	1980	A26	242997	117148	50,357905	5,675720
Underpass	Mixed	Unknown	NEUFCHATEAU	1987	A4	225230	64199	49,885034	5,414278
Underpass	Mixed	Unknown	Welkenraedt	1965	A3	259621	149783	50,648932	5,917556
Underpass	Mixed	Unknown	Welkenraedt	1965	A3	261272	149961	50,650219	5,940948
Underpass	Specific	BUSE 7 "Passage à gibier"	LIBRAMONT-CHEVIGNY	1984	A4	217607	66755	49,908351	5,309953
Underpass	Specific	Passage a gibier 1 (aérodrome)	SAINT-HUBERT	1900	N89	224080	83115	50,054651	5,403233
Underpass	Specific	Passage a gibier 2 (centre pénitenciaire)	SAINT-HUBERT	1980	N89	223308	81945	50,044229	5,392228
Underpass	Specific	Buse pour gibier - km 28.996	STOUMONT	1984	A26	245119	123449	50,414204	5,707133
Underpass	Specific	Buse pour bétail - km 30.239	STOUMONT	1985	A26	244924	122230	50,403278	5,704081
Underpass	Specific	Passage a gibier p192	TENNEVILLE	1980	N89	226673	86257	50,082565	5,440071
Underpass	Specific	Passage a gibier (ri de Bailet)	TENNEVILLE	1980	N89	224877	84061	50,063054	5,414547

1: Belge Lambert 72 projected system. 2: WGS84 geographic system.

Aerial view of the four overpasses build in Wallonia (specific passes above, mixed passes bellow).

Overpass for wildlife at Rulles

Overpass for wildlife at Grunhaut

Overpass for wildlife at Les Tailles

Overpass for wildlife at Baraque de Fraiture

APPENDIX 2: MACRO CODE FOR EXCEL

Important:

This program only works if IDs are recorded in the first column; if observations' dates are in the 8th column; and if X coordinates are in the 11th column and Y coordinates in the 12th column.
There must be a unique ID for each row. ID can be text or numerical.
Observations dates must have a date format in Excel.
X and Y coordinates must be projected coordinates (not latitude and longitude). One unit must correspond to one meter.

```
'*****************************************************************************
Option Explicit

Dim myA_Line_Counter As Integer
Dim myA_Identifier As String
Dim myA_DateObs As Date
Dim myA_LambertX As Long
Dim myA_LambertY As Long

Dim myB_Line_Counter As Integer
Dim myB_Identifier As String
Dim myB_DateObs As Date
Dim myB_LambertX As Long
Dim myB_LambertY As Long

Dim myAB_Distance As Long
Const myAB_Rayon_Limite As Integer = 500 'metres
Const myAB_Delai_Limite As Integer = 14 'jours
'*****************************************************************************

Sub TFE_Recherche_Doublons()

Application.ScreenUpdating = False

'insérer nouvelle colonne B pour les doublons
Columns("B:B").Select
With Selection
   .Insert Shift:=xlToRight
End With
Range("B1").Select
   ActiveCell = "Doublons"

'formater la nouvelle colonne B comme texte avec retour à la ligne
Columns("B:B").Select
Selection.NumberFormat = "@"
With Selection
   .HorizontalAlignment = xlLeft
   .VerticalAlignment = xlTop
   .WrapText = True
   .Orientation = 0
   .AddIndent = False
   .IndentLevel = 0
   .ShrinkToFit = False
   .ReadingOrder = xlContext
   .MergeCells = False
End With

'trier la feuille en ordre croissant sur la colonne A des Identifiants
Columns("A:A").Select
   ActiveSheet.Sort.SortFields.Clear
   ActiveSheet.Sort.SortFields.Add Key:=Range("A1"), _
      SortOn:=xlSortOnValues, Order:=xlAscending, DataOption:=xlSortNormal
   With ActiveSheet.Sort
```

```
        .SetRange Range("A1:AL10000")
        .Header = xlYes
        .MatchCase = False
        .Orientation = xlTopToBottom
        .SortMethod = xlPinYin
        .Apply
    End With

'pour chaque identifiant A en colonne A
'càd tant que l'identifiant A n'est pas = ""
'on démarre à la valeur 2 puisque la ligne 1 sont les titres
myA_Line_Counter = 2
Do Until Cells(myA_Line_Counter, 1) Like ""
    'sauver les données relatives à l'identifiant A
    myA_Identifier = Cells(myA_Line_Counter, 1)
    myA_DateObs = Cells(myA_Line_Counter, 8)
    myA_LambertX = Cells(myA_Line_Counter, 11)
    myA_LambertY = Cells(myA_Line_Counter, 12)
    'rechercher les identifiants B suivants qui seraient un doublon de l'identifiant A
    'càd tant que l'identifiant B n'est pas = ""
    'on démarre à la valeur myA_Line_Counter + 1 pcq on cherche dans les identifiants suivants
    myB_Line_Counter = myA_Line_Counter + 1
    Do Until Cells(myB_Line_Counter, 1) Like ""
        'sauver les données relatives à l'identifiant B
        myB_Identifier = Cells(myB_Line_Counter, 1)
        myB_DateObs = Cells(myB_Line_Counter, 8)
        myB_LambertX = Cells(myB_Line_Counter, 11)
        myB_LambertY = Cells(myB_Line_Counter, 12)
        'calculer la distance qui sépare les identifiants A et B
        'on utilise les coordonnées Lambert (X,Y) des identifiants A et B
        '= hypoténuse du triangle rectangle dont A et B sont les sommets non rectangles
        '  |2 /------------------
        '  | /      2        2
        '  |/ (Xb-Xa) + (Yb-Ya)
        '
        myAB_Distance = Sqr((myB_LambertX - myA_LambertX) ^ 2 + (myB_LambertY - myA_LambertY) ^ 2)
        'si la distance qui sépare les identifiants A et B est inférieure ou = au rayon limite et
        'si la différence de dates entre les identifiants A et B est inférieure ou = au délai limite
        'alors c'est un couple "doublon" potentiel et on indique
        '   l'identifiant A en regard de l'identifiant B en colonne B "Doublon"
        '   l'identifiant B en regard de l'identifiant A en colonne B "Doublon"
        If myAB_Distance <= myAB_Rayon_Limite _
            And _
            Abs(myB_DateObs - myA_DateObs) <= myAB_Delai_Limite Then
            If Cells(myA_Line_Counter, 2) <> "" Then
                Cells(myA_Line_Counter, 2) = Cells(myA_Line_Counter, 2) + "-" + myB_Identifier
            Else
                Cells(myA_Line_Counter, 2) = myB_Identifier
            End If
            If Cells(myB_Line_Counter, 2) <> "" Then
                Cells(myB_Line_Counter, 2) = Cells(myB_Line_Counter, 2) + "-" + myA_Identifier
            Else
                Cells(myB_Line_Counter, 2) = myA_Identifier
            End If
        End If
    'traiter l'identifiant "B" suivant
    myB_Line_Counter = myB_Line_Counter + 1
    Loop
'traiter l'identifiant "A" suivant
myA_Line_Counter = myA_Line_Counter + 1
Loop

'fin TFE_Recherche_Doublons
End Sub
'*****************************************************************************
```