Master Thesis

im Rahmen des
Universitätslehrganges „Geographical Information Science & Systems"
(UNIGIS MSc) am Interfakultären Fachbereich für GeoInformatik (Z_GIS)
der Paris Lodron-Universität Salzburg

zum Thema

„Automating the analysis of territory mapping data in bird monitoring"

vorgelegt von

Samuel Wechsler
104034, UNIGIS MSc Jahrgang 2015

Zur Erlangung des Grades
„Master of Science (Geographical Information Science & Systems) – MSc(GIS)"

Sempach, 31.8.2018

Eidesstattliche Erklärung

Ich erkläre hiermit eidesstattlich [durch meine eigenhändige Unterschrift], dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht. Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Bachelor-/ Master-/Diplomarbeit/ Dissertation eingereicht.

_____

Samuel Wechsler

Abstract

In an era of intense environmental change, monitoring biodiversity is a fundamental need. Bird populations have proven to be practical and meaningful indicators for tracking these changes. One of the key methods in monitoring bird populations is territory mapping. This work evaluated whether bird monitoring data collected with this method can be analysed automatically. The analysis of territory mapping data is not only time-consuming but also known to leave substantial scope for interpretation by human analysts. This subjectivity makes the analysis of territory mapping data susceptible to changes the way the data are interpreted across time and space. The call for automation is hence neither new nor surprising. But previous attempts to automate the analysis are either outdated (dating back to the 80's), failed to develop algorithms that were generally applicable for many species, failed to use methods that were flexible and adaptive to local conditions and input data or are only applicable to extremely labour-intensive forms of territory mapping (requiring 8-12 surveys per season). Two novel approaches were developed that estimate between-territory distances based on information in the recorded observation data. These site-, year- and species-specific distances are used to terminate a hierarchical clustering algorithm the right moment. Application on a large data set of the common breeding bird monitoring in Switzerland revealed that automatic analysis returns similar territory counts as manual solutions. Specifically automatic analysis was rather precise deviating from manual territory delimitation by only 12% on average. Globally the number of territories was overestimated by 2%. Comparison with accuracy of manual territory delimitation revealed that these deviations are of similar magnitude indicating that automation of analysis for many species is possible. Further optimisation of parameters is however recommended. Removal of subjectivity in the analysis of territory mapping by automation will reduce risks for biases in the data dramatically and is putting the analysis of already highly valued long term monitoring data on even more sound grounds.

Table of Contents

1 Introduction

1.1 Why is counting birds important?

Humans impact their environment and change it dramatically – climate change, deforestation or releases of pesticides like DDT are just some examples. These changes and their effect chains are often multivariate and complex, span across large spatial and temporal scales and are hence elusive for humans. In order to better understand and maybe halt or manage the impacts we have on our environment it can be helpful to reduce complexity by indicators that aggregate multiple aspects of environmental impacts (Bubb et al., 2010). The well-being of bird populations (especially population sizes) are considered to be good and practical indicators for the state of our environment (Bibby et al., 2000; Gregory and van Strien, 2010; Klvanova et al., 2009; Robinson and Sutherland, 2002). Birds seem prone to serve as indicators due to (at least) the following reasons:

- They inhabit almost all habitats and have diverse habitat requirements.
- They are often high up in the food chain and sensitive to change.
- They can live in close vicinity to humans, profit from their activities but also respond very sensibly to human-induced environmental impacts (e.g. habitat changes like agricultural intensification, pollutants like DDT).
- Compared to other taxonomic groups: They are relatively easy to count because they are well-known, easily recognisable and simpler to locate than many other organisms (Bibby et al., 2000). As a consequence many people are willing to volunteer in counting birds.
- For many countries there is, at least in Western Europe, good knowledge about population sizes and their development for decades – allowing for comparison with historical situations.
- Finally birds are also popular and are useful to convey high impact conservation messages.

According to Voříšek et al. (2008) collecting quantitative data on bird numbers serves two main tasks: status assessment and monitoring. Status assessment is used to evaluate the current situation and is an important prerequisite for prioritisation. The national share of a species international population is for example often an important criterion for initiating special protection measures at national level. Monitoring adds a temporal component by repeating counting events and hence allowing conclusions about trends and development of bird populations. This information can for example be used to measure success of conservation measures.

1.2 How to count birds

Even if it is easier than counting other organisms, counting birds is nevertheless not straightforward – there are challenges like avoiding double-counting or correcting for non-detected birds. Three commonly used methods are available: Line transects, point transects and territory mapping (Voříšek et al., 2008):

During line transects field workers follow a predefined line and record every bird seen or heard. Line transects are usually combined with distance sampling: Distances from the transect to each observation are recorded precisely or in distance bands (categories). Repeated visit help to cover the whole range of species from non-migratory birds to late-arriving trans-Saharan migrants. Line transects are a rather time-efficient method but it is sensitive to proper distance estimations (Bibby et al., 2000; Voříšek et al., 2008)

The point transect method is closely related to the line transect method and usually combined with distance sampling as well. Instead of constantly moving along a line transect, field workers record species and their distance from given points during a predefined duration without moving. Point counts are also considered time-efficient even though some time is "lost" moving from one point to the next. It is however quite sensible to errors in the distance estimates (more so than line transects) and not ideal for low-density or/and cryptic species that are missed easily during short, localised surveys (Bibby et al., 2000; Voříšek et al., 2008).

Territory mapping makes use of the territoriality of many bird species: During the breeding season individuals are often restricted to relatively small areas they actively defend against conspecifics and rarely leave in order to monopolize limited resources (i.e. nesting sites, food sources). For territory mapping bird observations are mapped in repeated visits inside a defined area. The idea behind the method is that accumulation of observations of birds' locations leads to clusters of observations that can be interpreted as territories(Voříšek et al., 2008). Mapping has the advantage of being spatially explicit allowing for additional analysis with respect to habitat besides the mere count data. However it is a rather time-intensive field method (Bibby et al., 2000) that takes a lot of time due to a high number of repeated visits and the complex analysis of field data after fieldwork (create territories from clusters). The analysis of the field data (i.e. the delimitation of the territories) is to some extent a subjective and demanding task and asks for the application of consistent rules and/or a thorough checking process (Voříšek et al., 2008). Additionally it is an inefficient method for non-territorial species, semi-colonial species, those that range wide, or those that are not monogamous (Bibby et al., 2000; Voříšek et al., 2008).

All three methods may correct statistically for imperfect detection (i.e. the fact that an observer never finds all birds actually present). In point and line transects the general idea is that detection probability is 100% at the position of the observer (the transect or the point) and becomes smaller with increasing distance. The relationship between distance and detection probability can be estimated with the data and used to correct for imperfect detection (Bibby et al., 2000; Voříšek et al., 2008). While for traditional territory mapping it was often assumed that all territories are found, in the last decades it has become common to also correct for imperfect detections in this method: In territory mapping detection probability can be modelled with the territory detection histories (or aggregations of it) in the course of repeated visits (Kéry et al., 2005; Kéry and Royle, 2016; Knaus et al., 2018; Royle et al., 2011, 2007, 2005).

1.3 Counting birds in Switzerland

Territory mapping has a long tradition in Switzerland and early forms were already in use by the late 1940s (Voříšek et al., 2008). It quickly became a standard and in 1993-1996 it was for the first time used on a broad scale for the 2nd Swiss breeding bird atlas (Schmid et al., 1998). Since 1999 it is in use in the common breeding bird monitoring scheme (MHB, "Monitoring häufige Brutvögel") where trends of common breeding birds are analysed with yearly surveys in 267 representative sample sites across Switzerland (Schmid et al., 2001). The method was also used in the latest Swiss breeding bird atlas 2013-2016 where territories in 2318 sample sites were mapped (Knaus et al., 2018).

1.3.1 The methodology in a nutshell

While traditional territory mapping requires up to 10 visits (Voříšek et al., 2008) in Switzerland a "simplified" territory mapping method with only 2-3 repeated visits per season is applied in most projects. Usually sampling sites of 1x1 km are visited on 2-3 mornings distributed regularly across the early

breeding season (Schmid et al., 2001). During these visits all bird observations are recorded on paper maps while following a predefined transect (left on figure 1). The observations are then digitized and for each species observations of all three visits are brought together on so-called "species maps" (right on figure 1). In a final step observations of the three visits are grouped into territories (transparently blue shaded polygons in figure 1). Important criteria for grouping observations are (Knaus and Schmid, 2014; Schmid, 2017; Schmid and Spiess, 2008):

- Always separate observations that are marked with a simultaneous observations.
- Try to have only one observation from the same visit per territory unless they are sexually compatible (i.e. two individuals that can reproduce together).
- Try to have observations from more than one visit per territory.
- Territory size varies naturally from species to species (e.g. Golden Eagle vs. Goldcrest). Territory sizes that correspond to the biology of a species should be applied when delimiting territories.

The number of territories per sample site is the desired result of simplified territory mapping while shape and location of territories are not relevant! The count data are then available for further analysis in order to correct for observer limitations like imperfect detection. Two prominent products in Switzerland that are based on territory mapping data are national species trends (Sattler et al., 2017) and abundance maps in atlas projects (Knaus et al., 2018; Schmid et al., 1998).
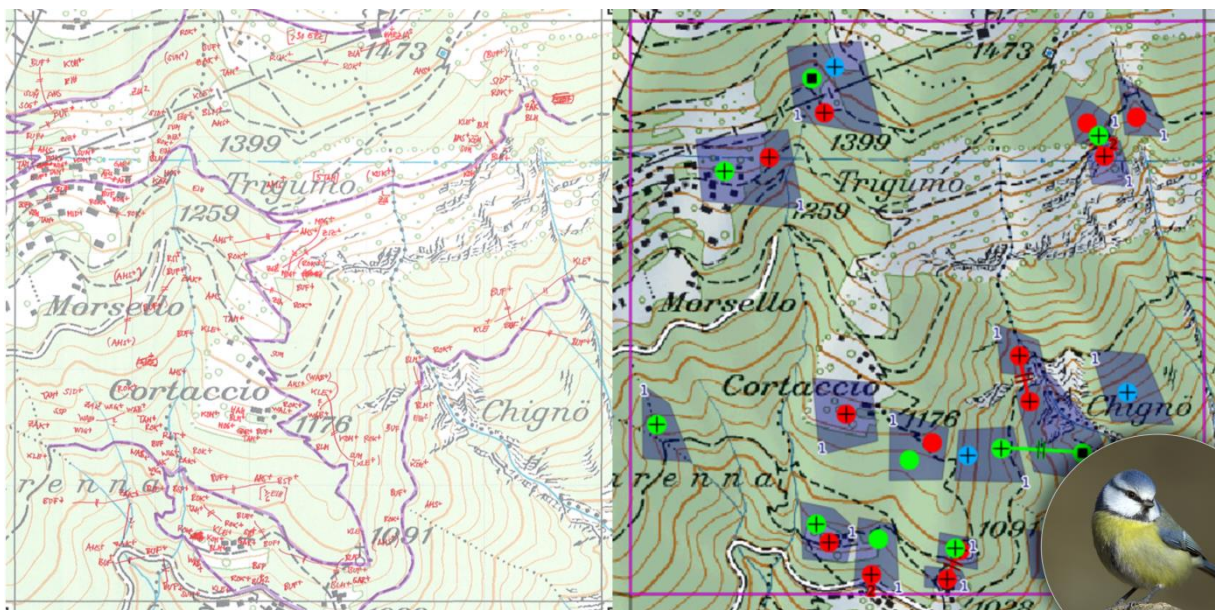
Figure 1: Territory mapping in progress. Left: Field map of one (out of three) visits at one example sample site (1x1 km square). Each observation is recorded with a three-letter-code for the bird species (all species of the visits are depicted). On the right the so-called "species map" for the Blue Tit *Cyanistes caeruleus* is shown. Different colours of the observations (points) represent the three different visits. Shaded polygons visualise the grouping of the observations to "territories". Base maps by swisstopo.

1.3.2 Important terms & expressions

In the course of this thesis some specific terms, linked to territory mapping in Switzerland, are used repeatedly. They are subsequently introduced and explained in detail. During territory mapping field workers are instructed to use a set of limited symbols to code the behaviour of each bird observation and also interactions to other birds. Table 1 presents an overview across all symbols and how they translate to other codes. The following expressions are the ones most crucial for the understanding of the thesis:

Simultaneous observation: Observations of two individuals at the same time (i.e. simultaneously) that indicate the presence of two territories. The most common example: Two simultaneously singing males. Such observations should always, by definition, be placed into two separate territories.

Double observation: Observations of the same individual at different locations. If a male is for example seen singing on two different roof tops this can be recorded with two observations linked by a double observation.

Potential double observation: Two observations that most likely but not certainly concern the same individual. During field work it can occur that an observation is made where circumstances indicate that it could well be a previously recorded individual. Such observations can be marked as potential double observations.

Flight observation: Observations of flying birds. Often the starting and/or the end point of the flight are not known. This is common for soaring birds, swallows and swifts but also for species that fly long distances to access ideal feeding grounds (e.g. Common Starlings *Sturnus vulgaris*).

Atlas code: Atlas codes are coding the behaviour of a bird (table 2). The purpose of these codes is to assess the probability that a species is breeding at a particular site. These codes were developed for atlas projects on breeding birds – hence the name. They are relevant to the algorithm because there are a species-specific minimal atlas codes that have to be met in order to create a valid territories (for detail see Knaus et al., 2018). In other words: Territories are only accepted if there is a certain likelihood of breeding.

Territory: Ecologically the term is referring to the space that a pair is occupying and defending against conspecifics during a breeding season in order to monopolize limited resources (i.e. nesting sites, food sources). In the context of simplified territory mapping "territory" is equivalent to a cluster/group of observations. While observations are grouped by means of polygons the shape of the polygon has no ecological meaning but only serves the purpose of grouping.

Table 1: Symbols used in territory mapping (digitized symbols), their meaning as well as their corresponding point/line code and atlas codes (table 2). Point and line codes are only relevant to readers that study the program code carefully. They identify the respective symbol.

Symbol	Meaning	Corresponding point/line code	Corresponding atlas code
♂	Male	14	2
♀	Female	15	2
⚥	Pair	16	4
+	Singing male present (or breeding calls heard) in breeding season	12	3
●	Other observation (seen, calling)	13	2
■	Imprecise localisation	2	2
⚡	Agitated behaviour or anxiety calls from adults	23	8
N	Transport of nesting material, nest-building or excavating of nest-hole by adult	11	10

F	Adult carrying food for young	17	16
J	Recently fledged young	22	13
○	Nest with breeding adult, eggs or young	19	18
	Simultaneous observation	3	-
	Potential double observation	9	-
	Double observation	8	-
	Flight observation (in the direction of the arrow)	4	2

Table 2: Atlas codes and definitions used in Switzerland (slightly modified from international standards).

Code	Meaning
	Possible breeding
1	Species observed during breeding period but suspected to be still on migration or to be a summering non-breeder
2	Species observed in breeding season in possible nesting habitat
3	Singing male(s) present (or breeding calls heard) in breeding season
	Probable breeding
4	Pair observed in suitable nesting habitat in breeding season
5	Permanent territory presumed through registration of territorial behaviour (song, etc.) on at least two different days a week or more apart at same place
6	Courtship and display (female and male)
7	Visiting probable nest-site
8	Agitated behaviour or anxiety calls from adults indicating presence of young or nest site close by
9	Brood patch on adult examined in the hand
10	Transport of nesting material, nest-building or excavating of nest-hole by adult
	Confirmed breeding
11	Distraction-display or injury-feigning
12	Used nest found (occupied or laid within period of survey)
13	Recently fledged young (nidicolous species) or downy young (nidifugous species)
14	Adults entering or leaving nest-site in circumstances indicating occupied nest (including high nests or nest holes, the contents of which cannot be seen) or adult seen incubating
15	Adult carrying a faecal sac of young
16	Adult carrying food for young
17	Eggshells of hatched young found (occupied or laid within period of survey)
18	Nest with breeding adult found
19	Nests containing eggs or young found

1.4 Need to automate

Territory mapping is long known two be (at least potentially) affected by two major sources of observer-related error (Best, 1975; Bibby et al., 2000; Enemar et al., 1978; Scheffer, 1987; Svensson, 1974; Verner and Milne, 1990):

1) Observational error due to different identification skills of observers, observation conditions, type of habitat (open country vs. forest) and conspicuousness of bird species

2) Interpretational error resulting from different interpretations of observational data (i.e. different delimitation of territories)

Observational error is generally believed to be of more importance or pose a greater threat to sound analysis of territory mapping data than interpretational bias (see for example Verner and Milne, 1990). Therefore it is not surprising that observational errors were at least partially tackled by means of statistical corrections for example for imperfect detection or incomplete coverage (Kéry et al., 2005; Royle et al., 2011). As this thesis is not focusing on this type of error, it will not be discussed further despite its importance. Instead we'll have a closer look at the second source of error that received far less attention.

The delimitation of territories is not an entirely objective task and leaves considerable scope for interpretation (Best, 1975; Bibby et al., 2000; Enemar et al., 1978; Scheffer, 1987; Svensson, 1974; Verner and Milne, 1990). While interpretational freedom may be beneficial in cases where regional knowledge can be incorporated in the delimitation process, it is however making territory mapping vulnerable to two biases:

1) Risk for spatial bias: If people in one corner of a study area delimit territories more generously than elsewhere spatial differences may appear that don't reflect reality. This would for example be problematic if abundance maps were modelled on biased territory mapping data.
2) Risk for temporal bias: Even if all sample sites were analysed by the same person and there was no spatial bias there remains a temporal uncertainty. The same analyst may produce different results when repeating the delimitation another year. If the delimitation became for example more generous over the years this may have a significant influence on species population trends.

To avoid such biases the "the application of consistent rules by all analysts" is required (Voříšek et al., 2008). Approaches to apply consistent rules are currently often limited to sound instructions and thorough checking of the analyst by experts. The Swiss Ornithological Institute is doing so by providing instructions (Schmid, 2017; Schmid and Spiess, 2008), individual feedback on conducted territory delimitations, example territory delimitations (Knaus and Schmid, 2014), yearly workshops for field workers and most importantly a rigorous checking of all territory delimitations by a handful of experts that remain the same for many years (see for example Knaus et al., 2018). But even with well trained and experienced experts variation in interpretations cannot be eliminated completely (see for example figure 2). Furthermore this checking and feedback process has its cost: The annual checking of 267 sample sites requires about 50 days of work (roughly 1.5h per site).

In order to apply rules consistently across time and space and save time in the analysis of territory mapping the call for automation is self-evident. In fact with the advent of computers first promising approaches to analyse territory mapping data have already been developed in England (North, 1977), Germany (Gerß, 1984) and the Netherlands (Scheffer, 1987). Of the 39 countries participating in the Pan-European Common Bird Monitoring Scheme (Klvanova et al., 2009) only Germany, Liechtenstein, Luxembourg, the Netherlands, Russia and Switzerland use territory mapping: Of those only the Netherlands are analysing since 2011 more than half of their sample sites automatically (van Dijk et al., 2013). For Germany and Switzerland who use both a simplified territory mapping approach with only 2-4 visits per site an automated approach is still lacking but highly desired.
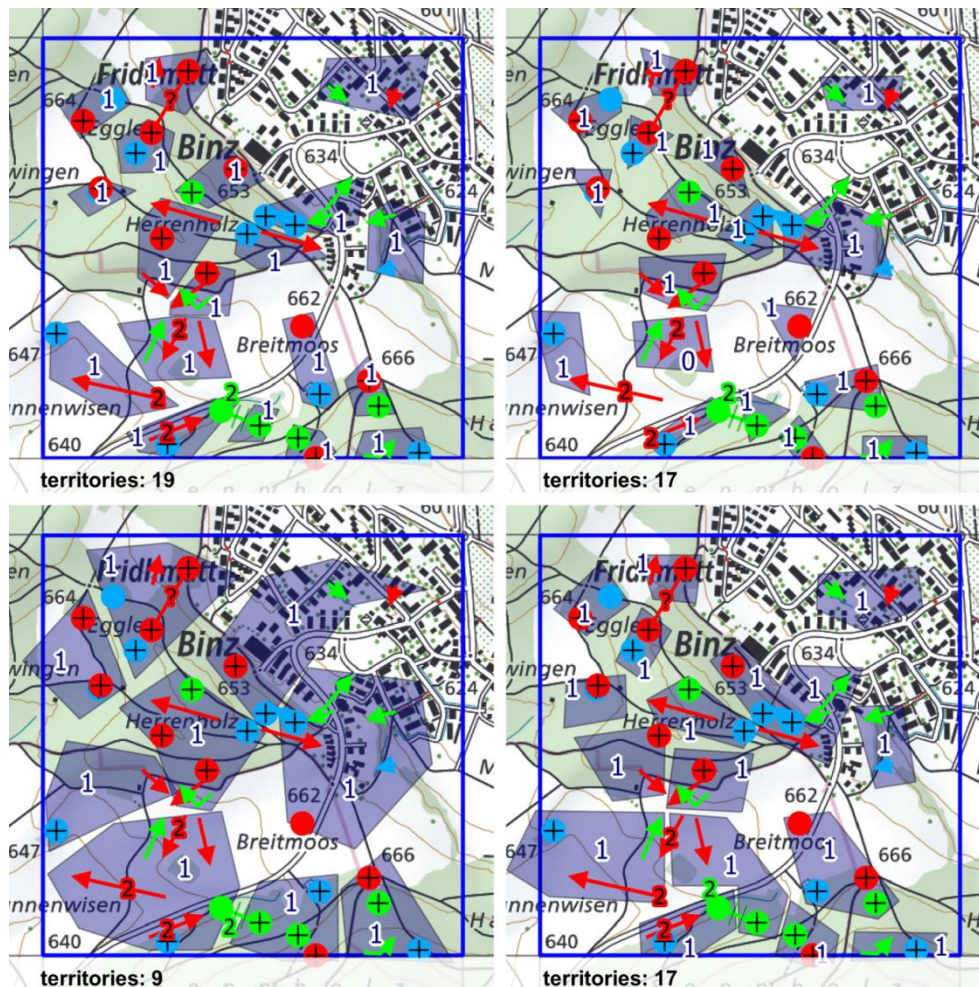
Figure 2: Four independent territory delimitations for the Common Wood Pigeon based on identical data. All four delimitations where made by experienced employees of the monitoring department of the Swiss Ornithological Institute. The total number of delimited territories varied between 9 and 19 territories. Variation illustrated in this square and species is large but not an extreme outlier. Base maps by swisstopo.

1.5 How to automate?

Despite the fact that territory mapping has been in use for a long time there are surprisingly few attempts to automate the analysis. While the shortcomings of manual territory delimitation are well known for some time (Best, 1975; Bibby et al., 2000; Enemar et al., 1978; Scheffer, 1987; Svensson, 1974; Verner and Milne, 1990) I only found five published algorithms that analyse territory mapping data.

1.5.1 Automatic analysis in existing algorithms

North (1977) developed an algorithm to analyse the British Common Birds Census data. Gerß (1984) developed an algorithm for German monitoring data. Scheffer (1987) and in 2011 van Dijk et al. (2013) developed solutions for the Netherlands. And finally Marchand (2015) published first ideas on how to analyse Swiss monitoring data. The only algorithm actively in use today is the one of van Dijk et al. (2013).

Common to these algorithms is that they all use hierarchical agglomerative clustering to group observations into territories. Cluster analysis seems indeed destined for such a task as it tries to "group

data objects based only on information found in the data that describe the objects and their relationships. The goal is that the objects within a group be similar to one another and different from the objects in other groups" (Tan et al., 2005). Hierarchical agglomerative clustering starts with as many clusters as there are data points and joins the two closest clusters step by step until all data points are joined in one big cluster. The second similarity of the five algorithms is that they all use Euclidian distance of projected coordinates (i.e. distance between observations in meters) as a dissimilarity or closeness measure.

There are however different methods to calculate what clusters are "closest". North (1977) and Gerß (1984) used the single linkage method that combines two cluster if the nearest points of existing clusters are closest. This method has the disadvantage of creating chain-like clusters. Scheffer (1987) improved this by using the average linkage method that defines "closeness" based on the distances between the centres of clusters. Marchand (2015) used Ward's method (Ward Jr, 1963) that is similar to the average linkage method as Ward's method assumes that a cluster is represented by its centroid, but it measures the proximity between to clusters in term of increase in SSE that results from merging the two clusters (Tan et al., 2005).

But already Scheffer (1987) wrote that the most difficult part in developing an automatism is defining the moment where the clustering process should stop. Completing clustering is leaving only one cluster/territory which is obviously not the desired result. North (1977) simply used one static maximal fusion distance above which points could not be combined and the clustering was stopped. This is unsatisfactory because this distance depends highly on the habitat and more importantly on the species analysed. Gerß (1984), Scheffer (1987) and Marchand (2015) assumed a sudden increase in the fusion distance the moment the algorithm moves from intra-territorial to inter-territorial distances and tried to find this point with various methods. However all these methods revealed occasionally unsatisfactory results.

The only operational method – the one of van Dijk et al. (2013) – is also using species-specific maximum distances. But their algorithm relies heavily on simultaneous observations (and observations from the same visit) so that these maximal distances are rarely used. Their algorithm tries to create as few clusters as possible but strictly prevents two observations from the same visit to be in the same territory. In other words: new clusters are created when points cannot be assigned to existing clusters because there are already observations of the same visit. Maximal distances only causes a new territory if an observation based on the "same visit" criterion could be grouped into an existing cluster but is farther away than the maximum distance. Their approach is based a lot on rules, i.e. whether or not two observations can be joined depending on the presence of simultaneous observations or observations from the same visit.

1.5.2 Alternative approaches?

It seems natural to use Euclidian distances of projected coordinates as "dissimilarity" input for clustering as we record bird observations in territory mapping on maps. Also Euclidian distance is a commonly used criterion to separate observations from different territories in manual analysis. However the distance metric could potentially also be of different type because pure linear distance (in meters) is not the only decisive criteria for territory delimitation. Close observation could nevertheless be from different territories if territories are separated by habitat types that act like barriers (e.g. highways or rivers). A potential approach to deal with such situations is calculating cost surfaces based on habitat variables (see for example Dean et al. 2015). It is for example imaginable that for some forest species crossing a deforested area is more "costly" (they don't like to do it) than moving inside a closed forest. This could be used for territory delimitation and implemented in computers with rasterized cost surfaces. For each cell a cost is defined. Based on this cost surface the least costly path/distance between any two points can be

calculated. The general approach is probably familiar to everyone that has ever used a route-finding algorithm to drive a car as fast as possible from A to B. For territory delimitation the cost surface could potentially combine different costs (e.g. slope, forest cover, roads or rivers). Instead of Euclidian distances least cost distances could potentially be biologically more meaningful. Nevertheless this path was not followed because of the following shortfalls: 1) Costs for crossing different habitat type have to be known, which is rarely the case (i.e. Is crossing a river of 10 m a problem for a Common Blackbird *Turdus merula*?). 2) For long term monitoring schemes cost surfaces of all sample sites would need to be updated regularly. 3) Habitat information, the basis for cost surfaces, is (currently) always outdated compared to bird observation data. 4) Cost surfaces and cell size need potentially be species-specific which would increase efforts for computation and data management (generate and update cost surfaces for 100+ species). 5) From personal experience the potential to improve clustering is doubtful as birds are highly mobile animals that can quickly cross unsuitable habitat. It's less the habitat that creates territory boundaries but interaction with conspecifics.

Principal Component Analysis (PCA) is another potential procedure to incorporate multiple characteristics (e.g. geographic distance and sexual distance) in order to find clusters in multivariate data. This approach was however not followed further because Principle component remains a black box that is difficult to understand or control.

Yet another approach to combine multiple characteristics is deep learning. It has been shown that using deep neural networks it is possible to transform multivariate data into more clustering-friendly representation (Aljalbout et al., 2018). This path was also not followed further because sufficiently understanding the complex theory behind deep neural networks was beyond the scope of this thesis and would potentially leave a black box that is similarly complicated to explain as principle component analysis.

1.5.3   Choice taken for this thesis

Based on the previous experience of other authors (specifically Marchand (2015)) and theoretical thoughts discussed above it was chosen to develop an algorithm that uses hierarchical agglomerative clustering with Euclidian distances as dissimilarity measures. Work in this thesis focused on finding the optimal moment to terminate clustering and incorporating non-Euclidian distance information (e.g. sexual incompatibilities) by modifying Euclidian distances.

1.6 Aims and research questions

The thesis pursued the following aims:

- Develop an algorithm that is capable of clustering observations from territory mapping, incorporating habitat, topographic and behavioural information.
- Compare clustering results (number of territories) with manual territory delimitation and it's variation.
- Support decision-makers on whether or not an automatic territory analysis could be implemented in long term-monitoring programs.

Based on these aims the following research questions were formulated:

1. Is it possible to produce one algorithm that gives reasonable results for all species of breeding birds?
2. Can an automatic analysis produce results that are comparable with results produced by manual territory delimitation?
3. Are there sensitivities to environmental conditions (e.g. area of settlements) or parameter values that could pose a risk for the long term use of the developed automatic solution?
4. Is calculation time a limitation in the use of such an algorithm?

2 Methods

2.1 About the raw data

The raw data used to develop the presented algorithm in this thesis consists of territory counts per sample sites and bird observations made during field work (e.g. the point and line symbols shown in figure 1). The data originate from the common breeding bird monitoring scheme (MHB = "Monitoring häufige Brutvögel") and the Swiss biodiversity monitoring (BDM) (Schmid et al., 2001). Both projects use the same methodology – simplified territory mapping in 1x1 km sample sites. The analysis incorporated data from 475 sample sites from the period 2013-2016 (figure 3). As most sites were surveyed annually this resulted in a total of 1268 surveys (each survey consisting of 2-3 visits).

Figure 3: Location of the 475 sample sites. Each square represent a sample site of 1x1km size. Relief by Institute of Cartography and Geoinformation, ETH Zurich.

2.2 Autoterri – an algorithm for automatic territory delimitation

In the course of this thesis an algorithm (actually three slightly different approaches) to automatically analyse territory mapping data was written. This chapter documents how this algorithm, named "Autoterri" works. Chapter 2.2.1 summarizes how the algorithms work and gives a brief overview. Chapters 2.2.2-2.2.8 then explain the algorithm and code parts in more detail.

2.2.1 How Autoterri works – summary of the most important concepts for quick readers

Autoterri groups bird observations into clusters named "territories". The clustering of observations is primarily based on Euclidian distances as similarity metrics: The closer two observations are the more likely that they are from the same territory. A second important concept applied is "compatibility". There are observations that are considered "incompatible" and independent of their Euclidian distance they should not (or less likely) be clustered into the same territory. Two singing males observed at the same time should for example not be grouped together. Compatibilities are used to modify (shorten or prolong) Euclidian distances. And thirdly the algorithm uses so-called "normalisation distances" that can be understood as "between-territory distances". If two observations are separated by more than the normalisation distance they will be grouped into two different territories. With these concepts in mind the general operating mode of Autoterri involves the following steps (explained in more detail in the indicated chapters):

1. Calculate Euclidian distances between all bird observations (chapter 2.2.3).
2. Modify these distances (shorten or prolong) according to their compatibility (chapter 2.2.4 and 2.2.6).
3. Perform hierarchical clustering on the modified distances using Ward's clustering method (Ward Jr, 1963) (chapter 2.2.7).
4. Cut the resulting dendrogram at the height of the species-specific normalisation distance. Observations that remain linked after the cutting are considered to be from the same territory. (chapter 2.2.5 and 2.2.7).

A very crucial aspect of this algorithm is the calculation of a reasonable normalisation distance because it is the deciding criteria whether two observations will be grouped or separated. Autoterri calculates this distance based on bird observations at a particular sample site in a particular year. The normalisation distance may hence vary between species, sites and years – depending on the distribution of observations. This is a desired property because it allows for example for higher territory density in bird-rich habitats compared to poor habitat. Subsequently I introduce the two most promising approaches for the calculation of the normalisation distance (out of three developed approaches):

The "**Quantile Value**" approach selects the normalisation distance from all distances that indicate between-territories distances with the help of a species-specific quantile. Let's elaborate this with an example: For the Common Chaffinch *Fringilla coelebs* the quantile value was set to q=0.035. So the 3.5% smallest distance of all the Euclidian distances that indicate different territories is used as the normalisation distance. Depending on the site and year this distance may for example be 86.3 m (random number chosen for illustrative purposes). So while the value of the quantile is static (but species-specific) the normalisation distance is not, as it depends on observational data that change from year to year and site to site.

The "**Mean Nearest Neighbour**" approach calculates a species-specific normalisation distance based on the average distance to the nearest neighbours. Distances used are filtered by their information content on the next neighbour (double observations are for example excluded because they only point to identical individuals but not to neighbours).

If there is not enough data to calculate meaningful normalisation distances (e.g. only two observations from different visits) or if calculated normalisation distances lay outside realistic minimum and maximum distance values, pre-defined and species-specific distance values are used as normalisation distances. For common species predefined distances are rarely applied.

While the previous paragraphs explain how Autoterri works in a generalised manner the following chapters will go more into the details of the algorithm. For this purpose the algorithm was divided into seven steps (figure 4).
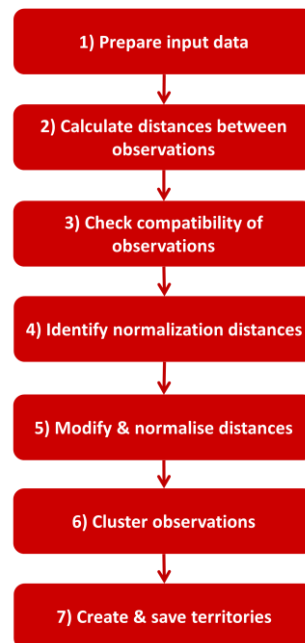
Figure 4: Overview over Autoterri

2.2.2 Step 1: Prepare input data

In the first part of the code all data necessary to continue with the algorithm is requested from a PostgreSQL database. This includes all observations (points and lines), as well as information regarding sex, number of individuals, altitude and whether two observations were marked as simultaneous, double or potential double observation in the field.

2.2.2.1 Extract observational data from PostgreSQL

Observational data are stored as points (observations) and lines (flights and relationships between observations) in a PostgreSQL database. The data request extracts these observations and also request the altitude for each observation form a digital elevation model in the data base (st_value). Point observations and flight observations (stored as lines) are saved to one table structure for the further course of the algorithm. For this purpose flights (lines) are represented by their centroid. Because flights may go outside the sample site only the part inside the sample site (st_intersection) is used for the calculation of the centroid. This guarantees that flight are always taken into consideration as soon as the bird once was inside the sample site. Also flights that are snapped to a starting point observation (for an example see figure 5) will be treated like double observations (i.e. starting point & centroid of flight = double observations).

Figure 5: Flight observation snapped to a starting point. Here an example for White-throated Dipper *Cinclus cinclus* singing and then flying downstream. Base map by swisstopo.

The following sql-statement illustrates exemplarily how flight observations are extracted. Point observations are extracted analogously. Point and line codes refer to the symbols introduced in table 3

```
req_sql <- paste("
select 'rel'||relations.id as id, relations.bird_id, linecodes.atlascode, round(st
_x(st_centroid(st_intersection(relations.geom,sampleareas.geom)))) as coord_x,
round(st_y(st_centroid(st_intersection(relations.geom,sampleareas.geom)))) as coor
d_y,
relations.linecode_id as pointcode_id, inventories.inventory_id, relations.count,
st_value(rast, st_centroid(st_intersection(relations.geom,sampleareas.geom))) as c
oord_z, linecodes.atlascode-birds.atlascode as ac_fullfilled, name_dt,
CASE WHEN to_date('",year,"-'||start_month||'-'||start_day, 'YYYY-MM-DD') > invent
ory_date
    THEN 1
    ELSE 0
    END as before_datelimit
from altitude, relations
INNER JOIN linecodes ON relations.linecode_id = linecodes.id
INNER JOIN inventories ON relations.inventory_id = inventories.id
INNER JOIN sampleareas ON sampleareas.id = inventories.samplearea_id
INNER JOIN birds ON relations.bird_id=birds.id
where relations.inventory_id in
  (select id from inventories where samplearea_id=(select id from sampleareas wher
e id
  =",samplearea,") and (extract (year from inventory_date) = ",year,"))",spchoice,
" and st_intersects(sampleareas.geom,
  relations.geom) and relations.linecode_id=4 and st_intersects(relations.geom, al
titude.rast) and count> 0 and st_value(rast, st_centroid(st_intersection(relations
.geom,sampleareas.geom))) is not null --centroid is important, otherwise lines tha
t cross 2 raster appear twice
order by coord_y desc", sep="")

req_flight <- dbSendQuery(con,req_sql)
df_flight <- fetch(req_flight,n=-1)
```

Besides the actual bird observations field workers also record if two observations somehow have a special relationship (i.e. if they are simultaneous, double or potential double observations). These data are also requested in this step and used later to infer for example incompatibilities.

2.2.2.2 Translate the count value into count-of-territory-values

The count value stored in the observation (points and flights) is used in the algorithm to increase the default territory count of 1 to higher values if multiple individuals were counted within 1 digitized observation. Hence the number of individuals needs translation into the number of territories that would correspond to such a number of individuals. The symbols recorded during field work are translated to number of territories (count_terri) according to the table 3. Two singing males correspond to two territories, while two warning adults will only lead to one territory.

Table 3: Field symbols and translation of recorded count values to count of territory values.

Symbol	Meaning	count	corresponding territory count
♂	Male	2	2
♀	Female	2	2
⚥	Pair (*take note that the count is by default set to 2 for this symbol*)	2	1
✚	Singing male present (or breeding calls heard) in breeding season	2	2
●	Other observation (seen, calling)	2	1
■	Imprecise localisation	2	1
⚡	Agitated behaviour or anxiety calls from adults	2	1
N	Transport of nesting material, nest-building or excavating of nest-hole by adult	2	1
F	Adult carrying food for young	2	1
J	Recently fledged young	10	1
◯	Nest with breeding adult, eggs or young	2	2
↗	Flight observation (in the direction of the arrow)	2	1

Some special cases were implemented as well. For the Red Crossbill *Loxia curvirostra* for example manual territory delimitation follows the rule of thumb that for every 5 individuals there is 1 territory (see Schmid and Spiess, 2008).

2.2.3 Step 2: Calculate distances between observations

Based on the data extracted from the PostgreSQL-DB in the step 1, different distances are calculated in step 2. These distances contain important information for modifications of Euclidian distances in step 5 and the clustering of observations in step 6.

- **Euclidian distance**: Distance in meters between any two points. Euclidian distance is calculated based on x-, y- and also z (i.e. altitude) coordinates as illustrated by the following code chunk: To my knowledge Autoterri is the first territory mapping algorithm that incorporates altitude for

distance calculation. A characteristic not irrelevant in areas with steep terrain – a large rock cliff for example may well separate bird territories on top and at the bottom of the cliff.

```
# creates a matrix with Euclidian distances from x, y, and z coordinate
dist.geo <- dist(df_obs[,c(4,5,9)], method="euclidean")
```

- **Temporal distance**: Binary variable distinguishing if two observations are from the same visit (value=0) or from different visits (value=1). This distance is important for compatibility reflections as some observations are incompatible if they are from the same visit (i.e. two singing males). Observations from different visits are per se compatible.

2.2.4 Step 3: Check compatibility of observations

In this step compatibility of observations is checked. Depending of their compatibility, factors for modifying the Euclidian distances between observations are defined that will be used in step 5. Compatibility is checked with respect to sexual compatibility (e.g. two males are incompatible) and behavioural observations (simultaneous observations are not compatible, double observations are highly compatible, potential double observations have an increased compatibility and flight observations have an increased compatibility due to spatial uncertainty). Factors defined for modifying distances depending on the compatibility of observations in step 4 are static and identical for all species.

2.2.4.1 Sexual compatibility

Based on the symbol chosen during field work (see table 2) each observation is assigned to one of the following:

- pair (point code 16)
- male (14, 12)
- female (15)
- juvenile (22)
- adult (13, 2, 23, 11, 17, 4)
- nest (19)

Some combinations of these symbols are sexually not compatible. Table 4 shows what symbols should not be placed into the same territory (i.e. are sexually incompatible) if they are from the same visit. While incompatibility is absolute for nests that cannot move around during a visit the incompatibility for other observations should not be absolute, as birds are mobile and may accidentally be recorded twice. For this reason Euclidian distances between nests are increased 100 fold while distances between the other "incompatible" observations are only increased by factor 1.5 in step 5 (chapter 2.2.6). Euclidian distances between compatible observations are not modified.

Table 4: Compatibility of observations and modification factors chosen to prolong distances between incompatible observations (red). Green fields indicate combinations of observations that are compatible.

	male	female	juvenile	pair	adult	nest
male	1.5			1.5		
female		1.5		1.5		
juvenile						
pair	1.5	1.5		1.5	1.5	
adult				1.5		
nest						100

The choice to increase Euclidian distances between nests by factor 100 is arbitrary – it could also be set to 1000. It only needs to be sufficiently high in order to guarantee placement of two nests recorded during the same visit into different territories. In other words this factor leads to binary decisions (same territory vs. not same territory). On the other hand the choice to increase distances between the other incompatible observations by one third is more subtle and critical: The factor should increase the chance of two "incompatible" observations to be in different territories but not prevent grouping in the same territory in an absolute manner. The influence of this parameter on clustering results was therefore analysed with a simple sensitivity analysis (see chapter 3.5.1).

2.2.4.2 Simultaneous, double and potential double observations

Simultaneous, double and potential double observations contain very informative observational information gathered directly in the field. This information is also used to check compatibility of observations and define how Euclidian distances between such observations are modified (table 5):

Table 5: Modifications of Euclidian distances by Autoterri between observations that were marked as simultaneous, double or potential double observations.

Type of observation	Modification of Euclidian distance
Simultaneous observation	Increase distance by factor 100
Double observation	Set distance to 0
Potential double observation	Decrease distance by factor 0.8

If observations are linked with a simultaneous observation symbol the distance between these two observations will be increased 100-fold in step 5 making it almost impossible that these two observations can end up in the same territory. The choice of increasing the distance by factor 100 should not be critical because it is so high that it is equivalent to a binary decision (placing to points with this symbol always into different territories).

If there are double observations (i.e. same individual at two different locations) their Euclidian distance is set to 0. Remember that observations concerning a flight observation that is snapped to a point are treated like a double observation. The choice to reduce these distances to 0 is not crucial because they concern the same individual. By reducing the distance to 0 it is guaranteed that the two observations are correctly place into the same territory.

If two observations are linked with a potential double observation symbol (line with ?) they will be marked accordingly and their distance is reduced to 80% of the original Euclidian distance. The idea behind this factor is bringing observations that could potentially concern the same individual closer together than normal observations. Hence potential double observations are more likely to be grouped into the same territory. The choice for a factor to deal with potential double observation distances could be anywhere

between 0 and 1 (not including 0 and 1). Despite the fact that the current choice is based on personal experience and is hence to some extent arbitrary it is not a crucial choice because only 0.2% of all distances (i.e. distances between the same species and the same visit) concern potential double observation. While the factor can be relevant in single situations the overall performance of the algorithm is certainly not strongly influenced by this choice.

2.2.4.3   Flight observations

Flight observations are more difficult to evaluate then normal observations. Flights may temporarily go outside a territory (figure 6), for example for visiting a temporarily available food source (e.g. freshly mown field) and they are generally less informative about the location of a territory. Distances to flight observations are therefore not considered when calculating the normalisation distance (i.e. the between-territory distance). Additionally the algorithm tries to group flights more easily to already existing territories instead of assigning them to new ones. Distances between any observation and a flight are for this purpose reduced to 67% of the original Euclidian distance. Thanks to this decrease flights fall more easily into close by territories. Reasonable values for such a factor to treat flight observations could be anywhere between say 0.3 and 1. The actual choice of 0.67 is based on personal impression and some basic trial and errors. It is not expected that this factor can influence the overall performance of the algorithm a lot because these distances are relatively rare. Only 2.5% of all distances (i.e. distances between the same species and the same visit) concern distances to flight observations.
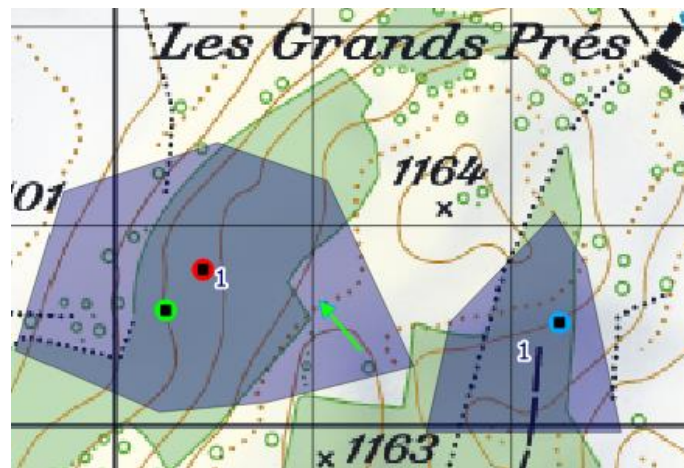
Figure 6: Showing a situation where a Common Wood Pigeon *Columba palumbus* flew outside of the core area of its territory (green arrow). Despite the distance to other observations it should be grouped together with the green and red point. Base map by swisstopo.

2.2.5   Step 4: Identification of normalization distance

The normalisation distance is a very crucial parameter of this algorithm. The normalisation distance can be understood as the minimal distance that separates two observations from different territories. If possible this parameter is calculated for each species and for each site and year separately based on the information contained in the observation data. This flexibility allows for between-year variation and also between-habitat variation (open grassland vs. dense alluvial forest). The normalisation distance is used in step 5 to normalise modified (due to compatibilities) Euclidian distances. Or in other words: The normalisation distance decides which (possibly modified) distances between points are two big to place points into the same territory. Three different approaches were tested in order to identify meaningful distances:

1) Linear Regression: Distances between all observations of a species originating from the same visit are used to create cumulative histograms of the distance measures. The values between the 25% and 75% quantile of the histogram are used for the calculation of a linear regression. The x-axis intercept of this linear regression is then used as the "normalisation distance".

2) Quantiles Values: In this approach distances between all points of a species are filtered to include only distances that point to distances between observations from different territories (i.e. exclude for example distances between males and females that can well be in the same territory). From the set of these distances a species-specific value (quantile) is defined in order to choose the normalisation distance - for example the 5% smallest distance.

3) Mean Nearest Neighbour: The idea of this approach is to calculate for each observation the distance to the closest neighbour of the same species that is most likely not from the same territory. The average of all these "nearest neighbour distances" is used as the normalisation distance.

2.2.5.1 Approach 1: Linear Regression

Marchand (2015) observed that cumulative histograms of distances between all observations contained smaller distances than cumulative histograms of only distances from the same visit. He concluded that the difference might be from the exclusion of within-territory distances in the histogram containing only the distances between observations from the save visit, assuming that the remaining distances contain information about the between-territory distances (i.e. the normalisation distance).

He proposed an approach where distances between all observations originating from the same visit are used to create cumulative histograms of the distance measures. In accordance to Marchand (2015) the values between the 25% and 75% quantile of the histogram containing all distances are used for a linear regression. The x-axis intercept of this linear regression is then used as the "normalisation distance" (see figure 7 for illustration of the approach).
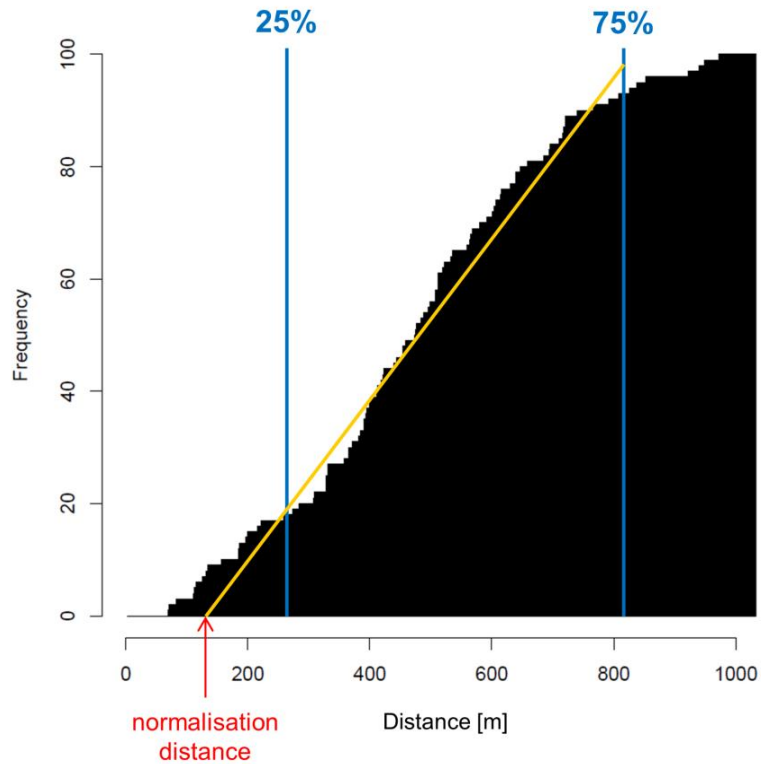
Figure 7: Schematic sketch of the Linear Regression approach. The cumulative histogram for a particular species shows distance measures between all observations selected. Blue lines indicate the 25% and 75% quantile. The yellow line depicts the linear regression. Modified after Marchand (2015).

2.2.5.2 Approach 2: Quantile Values

The distances between the observations contain a lot of information about the distance between territories. Nevertheless there are regularly some distances between observations that are smaller than the between-territory distance (e.g. if two rivals fight against each other at their territory borders). The idea behind this approach is to use the distances that seem "informative" but at the same time base the choice for the normalisation distance not on the smallest, most extreme value. Flight observations, double observations, potential double observations, distances between observations of different date and distances of observations that have no sexual incompatibility (e.g. female and male) are not considered for the calculation of the normalisation distance as these observations potentially contain no information about the distance between two territories (i.e. are not considered "informative"). From distances filtered this way the normalisation distance is chosen with the help of a species-specific quantile. The quantile-value is always bigger than 0 in order to not base the choice on the smallest available value. So if a species-specific quantile for species xy was for example set to 5% then the normalisation distance is equal to the 5% smallest distance of all the filtered distances. Figure 8 illustrates the general idea with the help of this example. So while the quantile is a static but species-specific value (i.e. 5%) the normalisation distance chosen is variable and depends on the distribution of the informative distances.
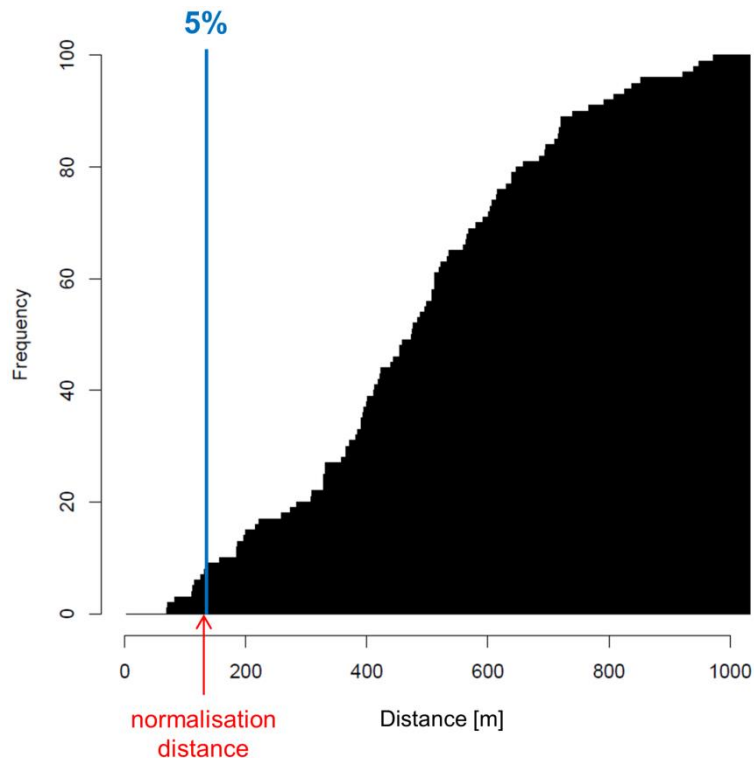
Figure 8: Schematic sketch of the Quantile Values approach. The cumulative histogram shows distance measures between all observations selected. The blue line indicates the 5% quantile that in this case is the value chosen for the normalisation distance that corresponds to about 170 m.

The following code illustrates how this approach is implemented.

```
# Loop species by species and calculate normalisation distance
for (spc in t_sp$species){

  # select distance data of observations that likely point to different
  territories
   selection<-sort(result$dist.geo[which(result$species==spc & result$doubleobs==0
   & result$potdoubleobs==0 & result$flight!=1 & result$dist.compatsex!=1 &
   result$dist.date==0)])

  # Calculate normalisation distance based on species-specific quantile of
  filtered distance measures and write normalization distance to result table
  q <- birds$quantile[birds$species==spc]
  result$dist.normalisation[which(result$species==spc)]<-quantile(selection, q)
}
```

As the following example code output shows the normalisation distance of the Common Chaffinch at this sample site was calculated to be 96.5 m which lies within the expert-defined ranges of 40 to 200 m. The value was found with the help of a quantile of 0.035.

```
## species          dist.normalisation  max_distance  min_distance   quantile
## Coal Tit                93.2             200            40           0.035
## Common Chaffinch        96.5             200            40           0.035
```

2.2.5.3   Approach 3: Mean Nearest Neighbour

This approach calculates a species-specific normalisation distance based on the average distance to the nearest neighbour (illustrated in figure 9). First, for each observation the nearest informative neighbour distance is found: Flight observations, double observations, potential double observations, distances between observations of different date and distances of observations that have no sexual incompatibility (e.g. female and male) are not considered as these observations potentially contain no information about the distance to the next neighbour and are hence considered "non-informative" (they could also be distances to the partner of the same territory). Then these distances are further restricted to lie between expert-defined minimum and maximum possible distance.
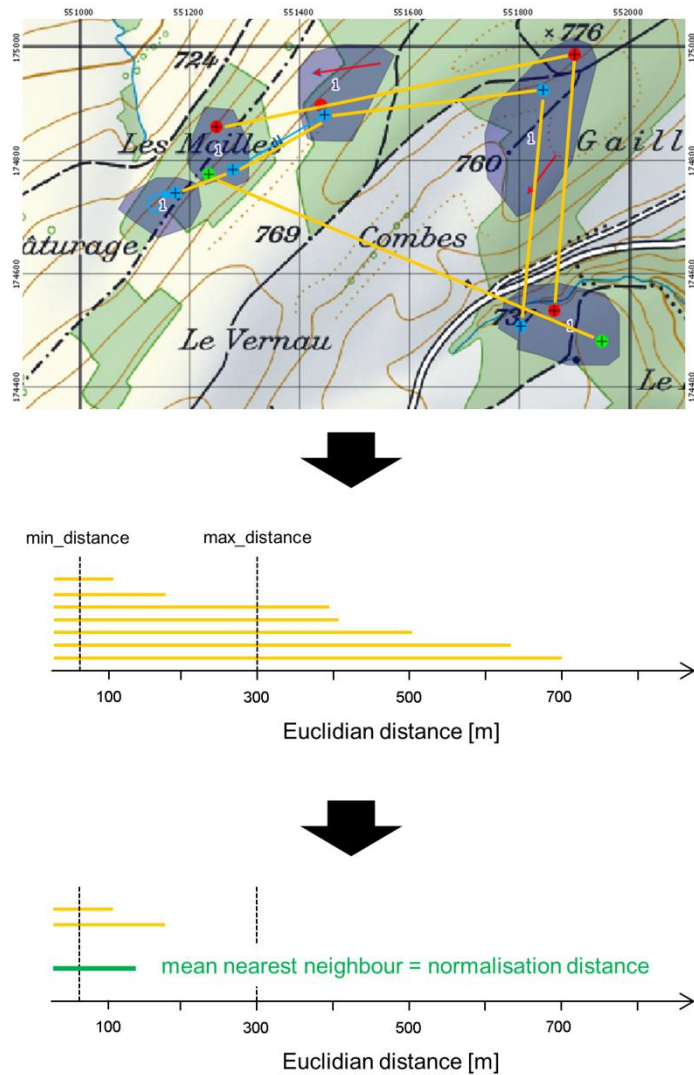
Figure 9: Mean Nearest Neighbour approach explained visually. First, informative nearest neighbour distances are identified (illustrated with yellow lines in the top illustration). Non-informative distances, i.e. between observations of different visits and compatible observations, are not considered. Secondly, these distances are limited to the expert-defined minimum and maximum distance (middle). Thirdly, the average of the remaining distances is calculated and used as the normalisation distance (bottom). Base map by swisstopo.

This restriction is important in situations where there are few observations and/or scattered breeding habitat. In these situations nearest neighbour distances can be artificially high and not representative for a value that indicates how closely neighbours could actually be (see figure 10 for exemplary illustration).

Figure 10: Example why limitation of distances between minimum and maximum distance is important for the calculation of the Mean Nearest Neighbour. Shown is an example for Eurasian Blackcap *Sylvia atricapilla*. All distances that indicate in principle nearest neighbour distances are indicated by lines. Red lines are above or below the maximum and minimum distance – and in fact they do not indicate meaningful nearest neighbour distances but distances that appear because suitable habitat (bushes) is very scattered in this sample site. Yellow lines however indicate true neighbours – i.e. distances that are probably established because of territorial behaviour of birds and not because of habitat distribution. Aerial image by swisstopo.

The following code snippet shows how this approach is implemented in R. Specifically the code shows how a matrix "nncomp" that contains information about whether a distance is informative with respect to between-territory distance is used to exclude non-informative distances (set to NA) from the mean nearest neighbour calculation. Then the mean nearest neighbour distance is calculated for each species with those distances values that are between the minimum and maximum distances.

```r
# calculate distances to nearest neighbours
nearest<-nn2(df_obs[,c(4,5,9)],df_obs[,c(4,5,9)], k=nrow(df_obs))
nn.idx<-as.data.frame(nearest$nn.idx)
nn.dist<-as.data.frame(nearest$nn.dist)

# Set distances that are non-informative (e.g. different species, sexually compa-
tible observations, double observations…) to NA
for (row in 1:nrow(nn.dist)) {
  for (column in 1:nrow(nn.dist)) {
    if (nncomp[row,nn.idx[row,column]]==0) {nn.dist[row, column]<-NA}
  }
}

# Extract the smallest distance/the nearest neighbour for each observation
nndistance<-apply(as.matrix(nn.dist), 1, min, na.rm=TRUE)
nndistance[nndistance==Inf]<-NA
```

```
# Loop species by species and calculate normalisation distance. Nearest neighbour
distances that are above or below the max_distance, min_distance are not
considered. They often concern distances of different corners of a sample site.
for (spc in unique(result$species[result$species!=0])){
  nndistance[which(df_obs$bird_id==spc &
  (nndistance>birds$max_distance[birds$species==spc] |
  nndistance<birds$min_distance[birds$species==spc]))]<-NA

  # Calculate the average smallest distance to the next neighbour
  avenn<-mean(nndistance[which(df_obs$bird_id==spc)], na.rm=TRUE)
```

2.2.5.4 Use of static normalisation distances when data-driven calculation of normalisation distance fails

If a normalisation distance cannot be calculated or the calculated distances is above or below a predefined, species-specific maximum or minimum values (mostly due to lack of sufficient data) the normalisation distance is replaced with a predefined maximum or minimum distance. The use of maximum and minimum distances is commonly applied in manual territory delimitation but distances are not explicitly stated and vary between different people. I defined these values for each species based on personal experience during correction of manual territory delimitation and the development of this algorithm. Examples for three species are given in table 6. How heavily the algorithm relies on this values is highly species-specific. For common species like the Common Chaffink 99% of the territories are delimited without using static distance measures. This value is lower in medium rare species like the Mistle Thrush *Turdus viscivorus* (58% in the Mean Nearest Neighbour approach) and even lower in rare species like the Eurasian Sparrowhawk *Accipiter nisus* (2%, Mean Nearest Neighbour approach). In the Mean Nearest Neighbour approach the value is not only used when no normalisation distance can be calculated but also to limit the values that go into the calculation for the Mean Nearest Neighbour. The value could be influential and leave room for species-specific adaptation/optimization of the algorithm. To evaluate the importance of these limits the influence of changing limits was elaborated for two species in chapter 3.5.2.

Table 6: Examples of species-specific values of the maximum and minimum normalisation distance.

species	max_distance	min_distance
Common Blackbird *Turdus merula*	200	40
Common Chiffchaff *Phylloscopus collybita*	200	50
Tree Pipit *Anthus trivialis*	300	50

2.2.6 Step 5: Modify & normalise distances

The general logic behind the algorithm is to run a hierarchical cluster analysis on modified and normalised Euclidian distances named "normalised distances". First, Euclidian distances between observations are normalized by the "normalisation distance". The later can be understood as the species-specific distance between territories ("Revierabstand" in German). By this normalisation step all species are made comparable/analysable together (the Common Kestrels *Falco tinnunculus* large distances are scaled down to the tiny distances of Common Firecrests *Regulus ignicapilla*). Secondly, these normalised distances are further modified by taking compatibilities into account, e.g. two singing males observed simultaneously should for example not be placed into the same territory even if their Euclidian distance is very small. Or,

to name another example of this 2nd step, distances between points that are linked by simultaneous observations are increased 100-fold so they are very unlikely to be placed into the same territory. The modifications in this 2nd step enlarge the normalised distances in case incompatibilities suggest so. They however also decrease the normalised distances in case it is indicated (for example for potential double observations). And finally the normalised distances between different species are set to the arbitrary high value of 5 (higher than the cut value of 1 used in step 6) to ensure that different species cannot be grouped into the same territory. The following formula summarises the modifications to the Euclidian distances:

$$normalised\ distance = \frac{\frac{Euclidian\ distance}{normalisation\ distance}}{compatiblity}$$

2.2.7   Step 6: Cluster observations

This part of the algorithm actually groups observation by creating a dendrogram (for an example see figure 11) using Ward's hierarchical clustering algorithm (Ward Jr, 1963). Hierarchical clustering starts with as many groups as elements and then continuously joins the 2 "closest" elements. This is repeated until all elements end up in one big cluster (Tan et al., 2005). In our situation the elements are bird observations. While many different methods for doing agglomerative hierarchical clustering are available, Ward's method seemed most promising when aiming at compact, regular clusters (Eger, 2018; Eisank and Koch, 2016; R Core Team, 2018): Ward's method attempts to minimize the sum of the squared distances of elements from their cluster centroids (Tan et al., 2005). Single and complete linkage methods were excluded on theoretical grounds because they are sensible to outliers and extreme points and/or are known to create chain-like clusters (Scheffer, 1987; Tan et al., 2005). The latter is an undesired behaviour in territory delimitation (Schmid and Spiess, 2008). Other available methods (average, centroid or median) were experimented with but did not reveal more convincing results.

Following the clustering the dendrogram is cut at a height of 1. The cut height can be set to 1 because all between-point distances were normalised with normalisation distances which represent the species-specific between-territory distances. This is equivalent to cutting the dendrogram at the species-specific normalisation distance when working with non-normalized distance measures. Observations that remain "linked" after the cut are grouped into the same territory.
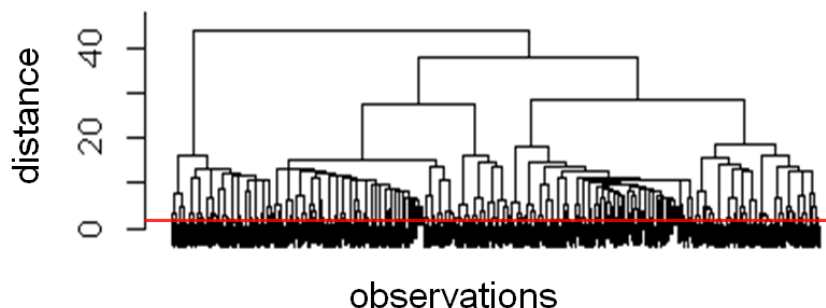
Figure 11: Dendrogram of observations at one sample site. Branches at the lowest level correspond to bird observations. These are continuously grouped until all observations belong to one single cluster. Cutting this tree at height=1 (red line) will leave some observations linked and hence members of the same territory. Distances used are "normalised distances" as explained in chapter 2.2.6.

The following code illustrates the hierarchical clustering with the function "hclust" (Müllner, 2013) and the subsequent grouping with "cutree".

```r
# Define "cut height"
cutheight <-1

# Creation of dendrogram
dendro<-hclust(result$dist.norm, method="ward.D2")
GRP <- cutree(dendro,h=cutheight)
```

2.2.8   Step 7: Create and save territories

In this final step territories in the form of polygons containing all the points of a particular cluster are generated. Points are grouped into a well-known-text multipoints and a buffer is created around these multipoints. Before saving the territories the validity of the resulting polygon is checked with respect to the following aspects:

- Minimal atlas code criteria: Territories are only accepted if they contain at least one observation that has an atlas code that is equal to or higher than the minimal atlas code criteria. For details on the species-specific atlas code criteria refer to Knaus et al. (2018).
- Date limit criteria: Territories are only accepted if they contain at least one observation after the species-specific date limit. The use of date limits is applied in territory mapping to exclude migrants. For details on the species-specific date limits refer to Knaus et al. (2018).
- Maximal altitude: Territories are only accepted if they contain observations that are below a predefined species-specific altitude. Observations above this altitude are not considered as "breeding birds". The same limits as used in the 3rd Swiss Breeding Bird atlas were applied (Knaus et al., 2018).
- Availability of breeding habitat: For some species territory validity is checked against the availability of suitable habitat. Specifically for Common Swift, Barn Swallow and Common House Martin territories are only accepted if there is a minimal area of buildings available in the sample site.

2.3 Quantifying errors in territory delimitation: measuring bias and precision

2.3.1   Requirements

In territory mapping the true number of territories is unknown. Known is however the number of territories delimited manually by the ornithologist who executed the fieldwork. This manual solution is always validated by experts of the Swiss Ornithological Institute and is hereafter referred to as the "original user's solution". All comparisons made in this thesis are made against the "original user's solution". It may seem unnatural to do so at first, because the validated solution of the original user could still be far away from the true territory number. In the absence of knowledge about the truth it is however important to guarantee continuity in the long-term data series of the common breeding bird monitoring. Any automatic solution should hence imitate what was done so far by hand as well as possible.

In order to evaluate the quality of Autoterri (or in fact any alternative territory delimitation) we are hence interested in knowing how far away we are from the original users (precision) and how much we generally over- or underestimate the number of territories (bias). Two statistical measures were developed in order to measure "accuracy" and "bias". While a "precision" measure should quantify the overall deviation from the manual territory count (i.e. the original user's solution), ignoring the direction of error (i.e. over- vs underestimating), a "bias" measure should quantify this systematic over- or underestimating but fail to detect a deviation if over- and underestimation compensate each other. Both are important features to evaluate alternative territory delimitations. In order to suit the nature of the here produced raw data such measures need to fulfil the following criteria. For explanatory purposes the criteria are illustrated giving examples referring to the fictive raw data table (table 7).

1. The absolute value of a measure should be **insensitive to the direction of the deviation** from the number of delimited territories by the original user, i.e. it should be symmetric. The measures should have the same absolute value if the automatic delimitation leads to a doubling or halving of the number of territories delimited manually. Example: Site 1 and 2 should lead to the same absolute values.

2. The measure should be **insensitive to sample size**. The measures value for the analysis of 2000 sites should be comparable with an analysis of only 100 sites. Example: The measure resulting from an analysis of only site 1 should be comparable with an analysis consisting of site 1 and 2.

3. The measure should be **weighted by the number of territories** involved. I.e. a deviation from 5 to 10 territories at one sample site should be less influential than a deviation of 10 to 20 territories at another sample site. Example: Site 3 should have more influence on a measure than site 2 because more territories are involved.

4. The measure needs to **be able to deal with 0 values**. The manual number of territories as well as the automatic number of territories can occasionally be 0. Example: Accuracy measures need to be computable also for sites like site 4 and 5.

5. The measure should be easily **interpretable**. The value should have some meaning that can easily be referred to the data.

Table 7: Fictive raw data for manually (original users solution) and automatically delimited territories for a fictive species "A" at 5 different sites. Ratio and difference of these values per sample site are also shown.

| site | species | Number of territories delimited | | Ratio | Difference |
		manually m	automatically a		
1	A	10	5	0.5	-5
2	A	5	10	2	+5
3	A	10	20	2	+10
4	A	2	0	0	-2
5	A	0	2	NA	+2

2.3.2 Mean Absolute Error (MAE) and Mean Error (ME)

Based on the above mentioned criteria the following measures were used to quantify the precision and bias in this thesis: Mean Absolute Error (MAE) for measuring the precision and Mean Error (ME) for the bias. Both measures are based on the comparison of the number of territories delimited manually m and automatically a as depicted in table 7. In its most basic form the measures can be calculated for the comparison of the two territory numbers at one sample site for one species (i.e. for one line in table 7). The values can be understood as the "average deviation per territory". To clarify the calculation the measures are below calculated exemplarily for sample site 1 of table 7:

$$MAE = = \frac{|a - m|}{\max(m, a)} = \frac{|5 - 10|}{\max(5, 10)} = \frac{5}{10} = 0.5$$

$$ME = = \frac{a - m}{\max(m, a)} = \frac{5 - 10}{\max(5, 10)} = \frac{-5}{10} = -0.5$$

This most basic form (i.e. comparison at one sample site for one species in one year) is aggregated in various forms to give information about precision and bias for entire species (not just at one sample site), different sample sites and the entire algorithm. The aggregation is achieved by taking weighted means of the MAE and ME values presented above. The values are weighted by the number of territories giving more weight to situations with many territories. Depending on the aggregation the MAE and ME are complemented with different subscripts to distinguish them: "sp" when values are aggregate across all sample sites for one particular species, "s" when values are aggregated across all species for one particular sample site and "algo" when values are aggregated across all species and sample sites for an overall measure of the algorithm.

Per species:

$$MAE_{sp} = \frac{\sum^s (\max(m, a) * MAE)}{\sum^s \max(m, a)} = \frac{\sum^s (|a - m|)}{\sum^s \max(m, a)}$$

$$ME_{sp} = \frac{\sum^s (\max(m, a) * ME)}{\sum_s \max(m, a)} = \frac{\sum_s (a - m)}{\sum_s \max(m, a)}$$

Per sample site:

$$MAE_s = \frac{\sum^{sp} (|a - m|)}{\sum^{sp} \max(m, a)}$$

$$ME_s = \frac{\sum^{sp} (a - m)}{\sum^{sp} \max(m, a)}$$

Per algorithm:

$$MAE_{algo} = \frac{\sum^{sp,s} (|a - m|)}{\sum^{sp,s} \max(m, a)}$$

$$ME_{algo} = \frac{\sum^{sp,s} (a - m)}{\sum^{sp,s} \max(m, a)}$$

m = count of manually delimited territories by original user of particular species sp and at particular site year combination s

a = count of automatically delimited territories of particular species sp and at particular site year combination s

s = year and sample site combination

sp = particular species

2.3.3 Comparison with manual territory delimitation

Manual territory delimitation is not an entirely objective task. Subjective decisions can create different outcomes (see for example figure 2). In order to evaluate the quality of any automatic territory delimitation it will be important to know the magnitude of error inherent to the manual territory delimitation to which results of an automated analysis will be compared to. In order to quantify this error repeated manual territory delimitation was conducted for 7 selected species in 35 representative sample sites independently by 8 different people (7 "experts" and 1 "original user").

2.3.3.1 Site selection

A stratified random sample site selection lead to 35 sample sites evenly distributed across the biogeographic sub-regions of Switzerland (Gonseth et al., 2001) as depicted in figure 3.
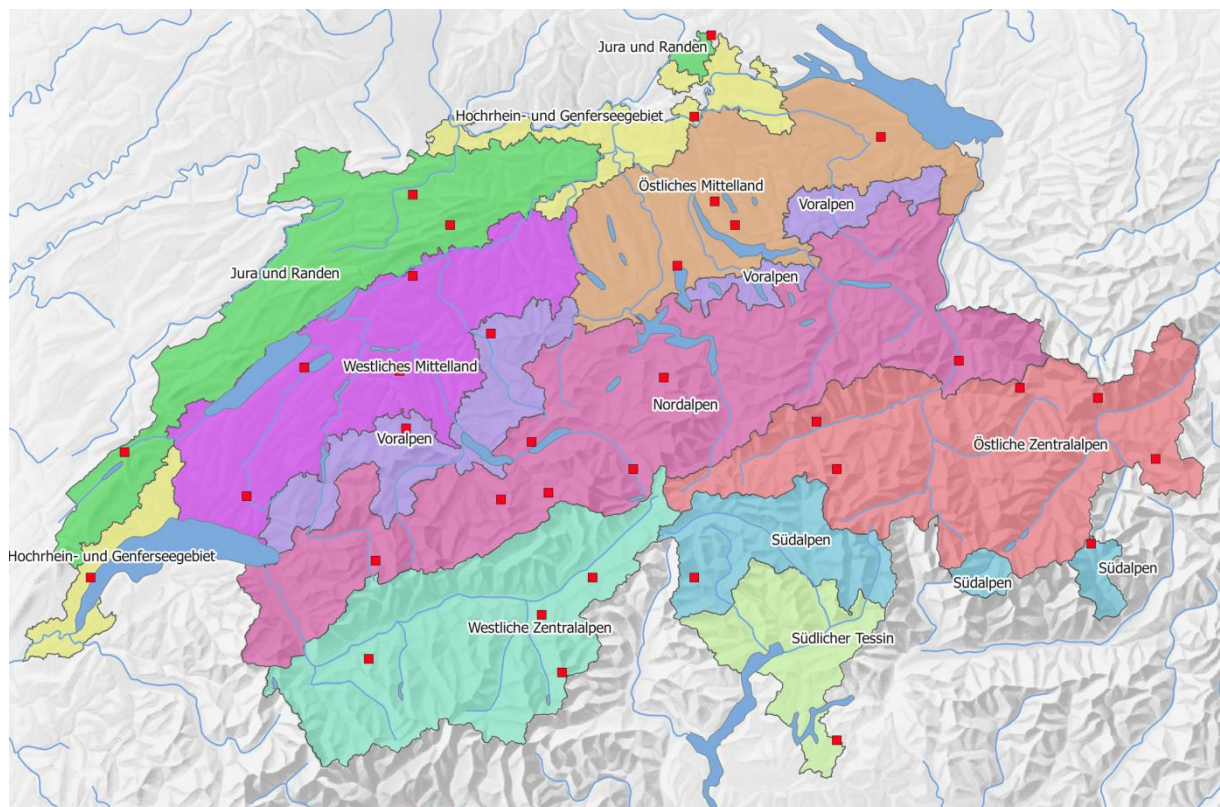
Figure 12: Biogeographic sub-regions of Switzerland and the location of the 35 sampling sites (red squares) used for error quantification of manual territory delimitation. Relief © Institute of Cartography and Geoinformation, ETH Zurich.

Apart from good biogeographical coverage sample sites were also chosen to be representative for the following six covariates because it is at least imaginable that they could influence the amount of uncertainty during manual territory delimitation:

- **Forest**: The percentage of forest cover within the 1x1 km sample site. Sites with dense forest are usually more difficult for manual territory delimitation because of higher bird numbers and poorer visibility/orientation.
- **Buildings**: The number of square meters covered by buildings within the 1x1 km sample site. Sites with very high share of buildings (dense cities) are usually more difficult for manual territory delimitation because of higher bird numbers and poorer visibility/orientation.

- **Farm**: Percentage of the area covered with farmland (meadows and acres). On Farmland it is thought to be easier to manually delimit territories due to lower bird density and better visibility/orientation.
- **Roads**: Length of all the roads inside the 1x1 km sample site. Usually easily accessible sample sites can be sampled with more spatial precision and hence manual territory delimitation is thought to be easier.
- **Elevation**: Average elevation of the sample site. Elevation is always a very important variable in Switzerland. It is often linked to vegetation coverage, bird densities, etc. and could potentially influence uncertainty in manual territory delimitation.
- **Route length**: The length of the path walked during the visits (i.e. during bird counting). It is highly likely that longer routes lead to better spatial coverage and more observations during the sampling process. This could also influence the amount of error during manual territory delimitation.

Figure 13 shows that the selected 35 sample sites were representative for the Swiss landscape (or in case of the "route length" for the Atlas sampling scheme) with respect to the 6 covariates.
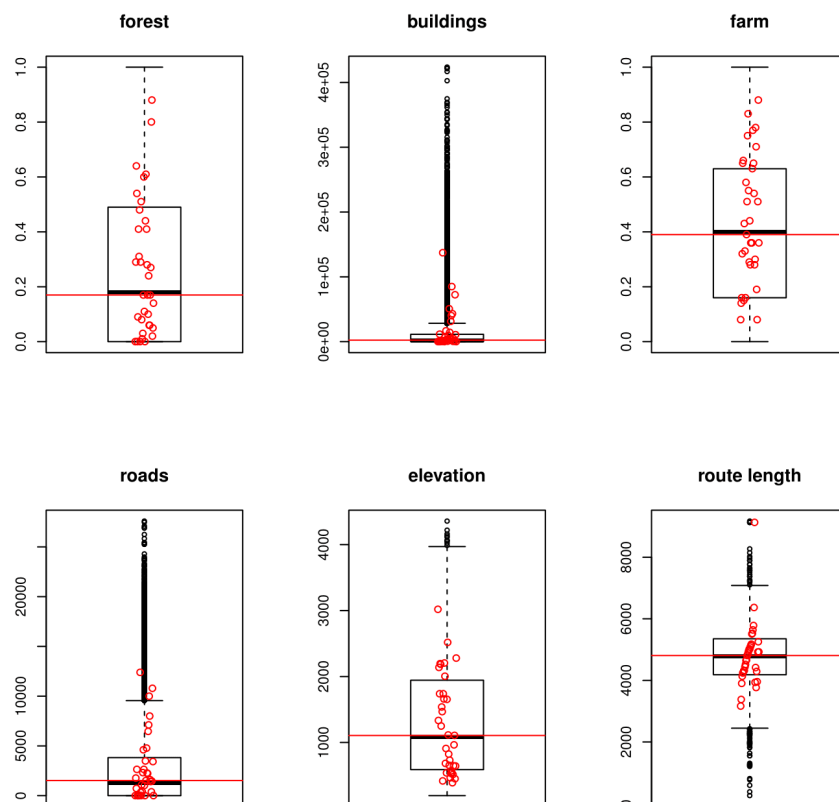
Figure 13: Sample sites in comparison to distributions in Switzerland. Boxplots showing (in black) the distribution of the variables: percentage covered with forest, m² of buildings, percentage covered with farmland, length of roads, average elevation of 1x1 km squares of the Swiss national grid (N=41301). The boxplot for route length shows the distribution of route length for the sampling sites used in the atlas 2013-2016 (N=2148). Red circles show the values for the 35 selected sampling sites. The red horizontal line shows the median value of the 35 selected sites.

2.3.3.2 Species selection

In order to keep the work load at a bearable level a choice on what species should be analysed had to be made. Based on the following criteria seven species were finally selected for the analysis (table 8).

- Species should be sufficiently common to be well represented and usually have more than 1 territory inside a particular sample site. If there is only one territory, delimitation is normally straight forward. In the set of 1268 sampling events conducted between 2013-2016 a possible species had to have at least 1000 manually delimited territories and on average >3 territories per site (not considering sites that lack the species) to be a potential candidate.
- The species should be of different systematic families in order to cover different ecological groups and hence behaviour and nesting strategies.
- The species list should include species that are expected to be easy to analyse as well as some expected to be difficult for an automated territory delimitation approach. From manual territory delimitation we know already by experience that species that clearly occupy and defend a territory are better suited for territory delimitation than species that fly vast distances in search of food or behave colonial or the like. This point was hence included in a more or less subjective manner based on preliminary results, experience during algorithm development and experience in manual territory delimitation.

Table 8: List of the seven species selected for error quantification in manual territory delimitation. Also the expected behaviour in (automated) territory delimitation is indicated.

Species	Expected behaviour
Common Wood Pigeon *Columba palumbus*	difficult
Black Redstart *Phoenicurus ochruros*	easy
Common Blackbird *Turdus merula*	easy
Eurasian Blackcap *Sylvia atricapilla*	easy
Goldcrest *Regulus regulus*	easy
Common Starling *Sturnus vulgaris*	difficult
European Greenfinch *Carduelis chloris*	difficult

2.3.3.3 Error quantification: Coefficient of Variation (CV), MAE & ME

In order to quantify the error in manual territory delimitations three different values were calculated: For comparison with automatic territory delimitation on the one hand mean absolute error (MAE_{sp}) and mean error (ME_{sp}) were calculated analogously as explained previously. Expert's solutions are hence set in relation to the original user's solution. In the formulas presented in chapter 2.3.2 automatic territory count values a are simply replaced by territory count values of the experts delimitations.

On the other hand studies quantifying the variance due to manual territory delimitation by different analysts often reported the coefficient of variation (CV_{sp}). For the sake of perfect comparability of errors due to manual delimitation with data from literature the coefficient of variation was also calculated and referenced in some chapters of this study. Figure 14 shows however that MAE_{sp} and CV_{sp} are highly correlated and hence comparable.

$$CV = \frac{\sum (\bar{m} * \frac{sd(m)}{\bar{m}})}{\sum \bar{m}} = \frac{\sum (sd(m))}{\sum \bar{m}}$$

$\bar{m}$ = average number of manually delimited territories (experts and original user) for a particular species and at particular site-year combination from manual territory delimitation by multiple analysts

sd(m) = standard deviation of the number of manually delimited territories (expert and original user) for a particular species and at particular site-year combination from manual territory delimitation by multiple analysts
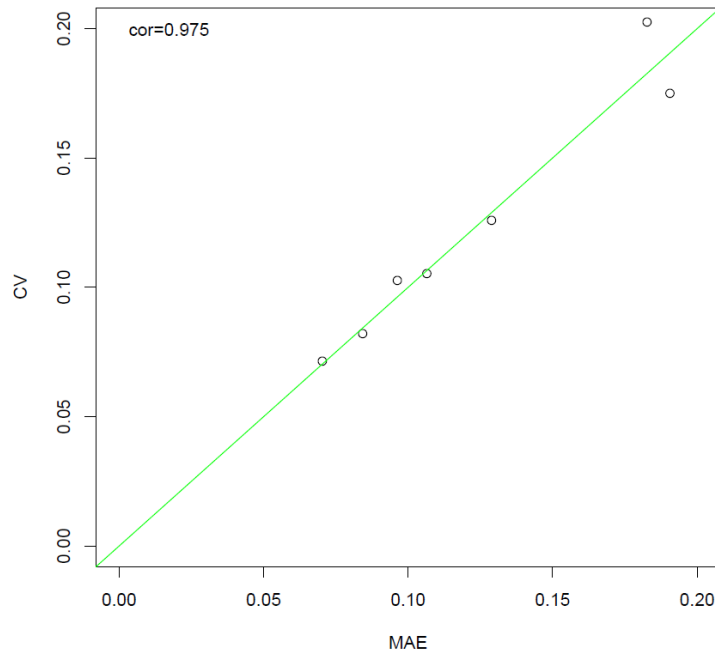
Figure 14: Comparison of coefficient of variation (CV) and mean absolute error ($\boldsymbol{MAE}_{sp}$) for seven species analysed by multiple experts. The green line corresponds to the ideal case of complete accordance of the two measures. Paersons correlation was 97.5%.

2.4 Optimizing quantiles for the Quantile Value approach

When identifying the normalization distance with the Quantile Value approach (see chapter 2.2.5.2) the species-specific quantile is a rather influential parameter. In an initial version of this approach this parameter was chosen for every species based on expert knowledge and some trial and error experience. To assess the potential to improve Autoterri this parameter was optimized for the data of all sample sites of the years 2013 and 2015. In a second step the improvement in MAE_{sp} and ME_{sp} was evaluated based on data not involved in the optimization process: Namely all sample sites for the years 2014 and 2016.

A first approach choosing quantiles randomly form a uniform distribution and repeating this 1000 times proofed unfeasible in the scope of this thesis due to very long computation times. Instead Autoterri (approach "Quantile Value") was run for each species and for 60 predefined quantiles ranging from 0.005 to 0.1 in steps of 0.005 and from 0.1 to 0.5 in steps of 0.01. During the optimization process the limitation of the normalisation distance between a species-specific minimum and maximum was deactivated so results were not masked by these limits. For each quantile territories were automatically delimited and accuracy measures calculated (figure 15). The best quantile was chosen based on the minimal value for the sum of $MAE_{sp}+ME_{sp}$. As ME_{sp} is by its nature smaller than MAE_{sp}, MAE_{sp} has more weight in this combined value. For 95 out of 163 species an optimum quantile was identified (see appendix table 21). For 67 species it was not possible to identify an optimised quantile mainly due to lack of informative data (figure 15). For many raptors for example there is usually only 1 territory per sample site which makes it impossible to calculate an optimal quantile.
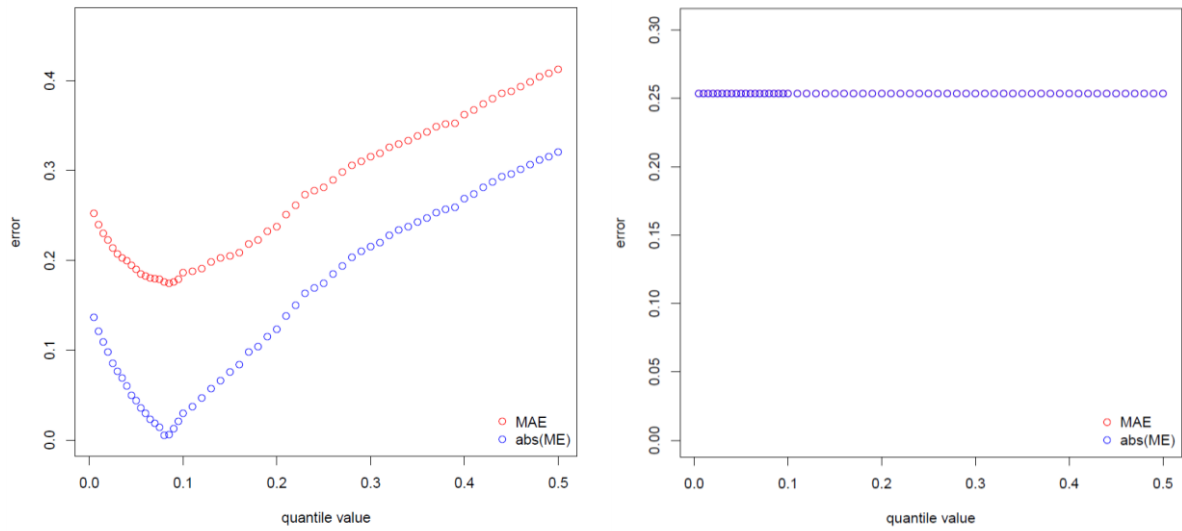
Figure 15. Accuracy ($\textbf{MAE}_{sp}$ and absolute values of $\textbf{ME}_{sp}$) for different quantiles show pronounced optima for many common species like the European Greenfinch (left) that seems to have a minimal deviation to the manual territory delimitation at a quantile value of 0.085. For 67 mainly rare species like the Eurasian Sparrowhawk (right) finding an optimum often failed because there is rarely more than 1 territory per sample site available to test the influence of different quantiles.

3 Results

3.1 Accuracy of Autoterri

3.1.1   Per algorithm

Autoterri was run with three approaches to calculate the normalisation distance. The overall deviation from the manual territory delimitation is summarized in table 9. Across all 162 species and all 475 sample sites in 2013-2016 the overall MAE_{algo} varied between 17.7% for the Linear Regression approach and 12.0% for the Mean Nearest Neighbour approach. The bias (ME_{algo}) varied between 4.8% for the Linear Regression approach and 2.1% for the Quantile Value approach. All approaches tend to overestimate the number of territories. Pearson's correlation between the number of manually delimited territories and the automatic solution at each species-year-site combination was always high – at maximum 0.973 for the Mean Nearest Neighbour approach.

Table 9: Accuracy measures for Autoterri run with three different approaches for the calculation of the normalisation distance. r_{algo} denotes the Pearson's correlation between the number of manually delimited territories and the automatic solution at each species-year-site combination.

Accuracy	Approach for identification of normalisation distance		
	Linear Regression	Quantile Value	Mean Nearest Neighbour
MAE_{algo}	0.177	0.123	0.120
ME_{algo}	0.048	0.021	0.023
r_{algo}	0.946	0.972	0.973

3.1.2   Per sample site

When aggregating the accuracy by sample site the three approaches compared as shown in figure 16. For the Linear Regression approach the median MAE_s equaled to 16.4% (min: 0%, max: 53.3%) and the median ME_{sp} was 4.7% (min: -37.9, max: 53.0%). For the optimised Quantile Value approach the median MAE_s equaled to 11.4% (min: 0%, max: 52.9%) and the median ME_{sp} was 1.8% (min: -36.8%, max: 46.9%). And for the Mean Nearest Neighbour approach the median MAE_s equaled to 11.0% (min: 0%, max: 49.2%) and the median ME_{sp} was 2.5% (min: -36.3%, max: 42.1%).
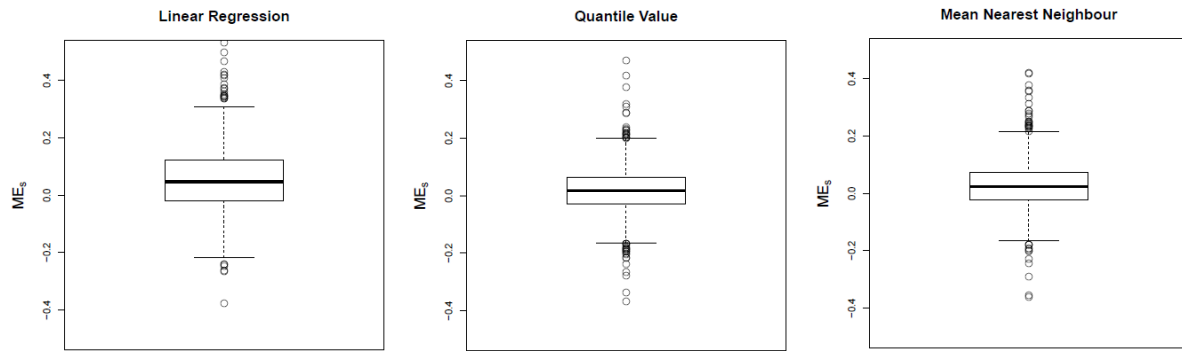
Figure 16: Distribution of ME_s for the three different approaches for calculating the normalisation distance.

Figure 17 illustrates exemplarily how manual and automatic solutions compared at three different sites. Sample site 505 had a low MAE_s of 1.7% while the MAE_s in sample site 408 was 31.7%. The accuracy of sample site 338 was with a MAE_s of 11.4% in between. Sample site 408 contains some dramatic outliers (species that produced completely different automatic than manual territory counts).
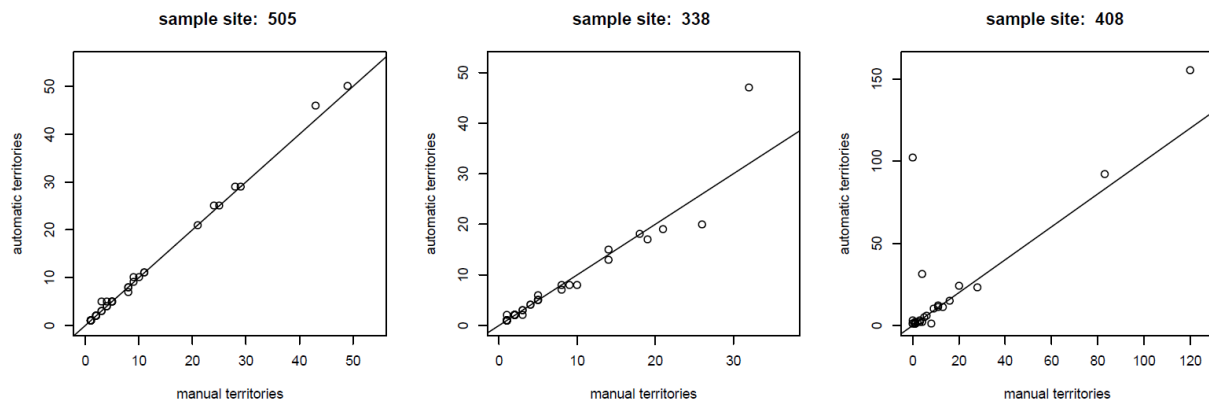
Figure 17: Manually delimited territories vs. automatically delimited territories at three different sample sites with low MAE_s (left) to high MAE_s (right). Shown are results analysed with the Mean Nearest Neighbour approach. Points represent the different species. The black line indicates the ideal case of perfect fit between manual and automatic territory delimitation. Points above the black line hence indicate an overestimation of the number of territories for a particular species compared to the manual solution of the original user.

3.1.3 Per species

Performance varied largely between species. Table 10 shows MAE_{sp} and ME_{sp} values for the 10 most common species (for all species refer to appendix 0). The Mean Nearest Neighbour approach overestimated the number of Common Chaffinch territories for example by 0.8% (ME_{sp}). In absolute numbers this corresponds to 34'633 automatically delimited vs. 34'349 manually delimited Common Chaffinch territories.

Table 10: Accuracy values for the 10 most common species. Values are given for the three different approaches for identifying the normalisation distance. The list is sorted according to number of the manually delimited territories analysed from 2013–2016. Scientific names for species not previously mentioned: European Robin *Erithacus rubecula*, Coal Tit *Periparus ater*, House Sparrow *Passer domesticus*, Eurasian Wren *Troglodytes troglodytes*, Great Tit *Parus major*, Song Thrush *Turdus philomelos*, Black Redstart *Phoenicurus ochruros*.

Species	manual count	Linear Regression MAE_{sp}	Quantile Value MAE_{sp}	Mean Nearest Neighbour MAE_{sp}	Linear Regression ME_{sp}	Quantile Value ME_{sp}	Mean Nearest Neighbour ME_{sp}
Common Chaffinch	34349	0.164	0.083	0.076	-0.075	0.004	0.008
Eurasian Blackcap	20734	0.153	0.078	0.071	-0.054	0.019	-0.017
Common Blackbird	19205	0.163	0.113	0.097	-0.007	0.018	0.007
European Robin	16134	0.145	0.093	0.085	0.008	0.018	-0.004
Coal Tit	15799	0.155	0.095	0.085	-0.028	0.015	-0.013
House Sparrow	15292	0.192	0.175	0.203	-0.105	-0.082	-0.131
Eurasian Wren	14378	0.155	0.084	0.076	0.035	0.019	-0.003
Great Tit	13596	0.149	0.112	0.101	-0.013	0.042	0.001
Song Thrush	9766	0.132	0.096	0.096	0.057	0.017	0.052
Black Redstart	9209	0.151	0.103	0.092	0.093	0.034	0.038

The plots in figure 18 show for the ten most common species how automatically delimited territory numbers per site compared to manually delimited territories. While these plots are very intuitive to read one has to keep in mind that exceptional values are visually more prominent because they are often less covered by results in other sample sites. Results of one dot in in figure 18, i.e. the territories delimited, are plotted exemplarily on maps in figure 19 (for the Common Chaffinch).
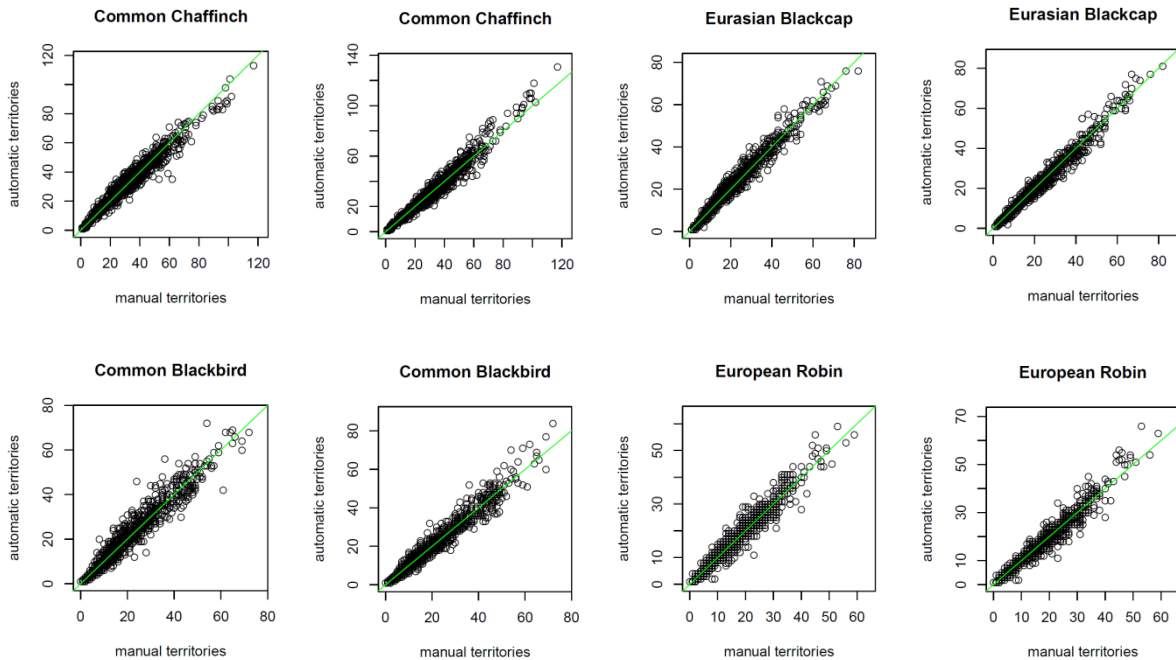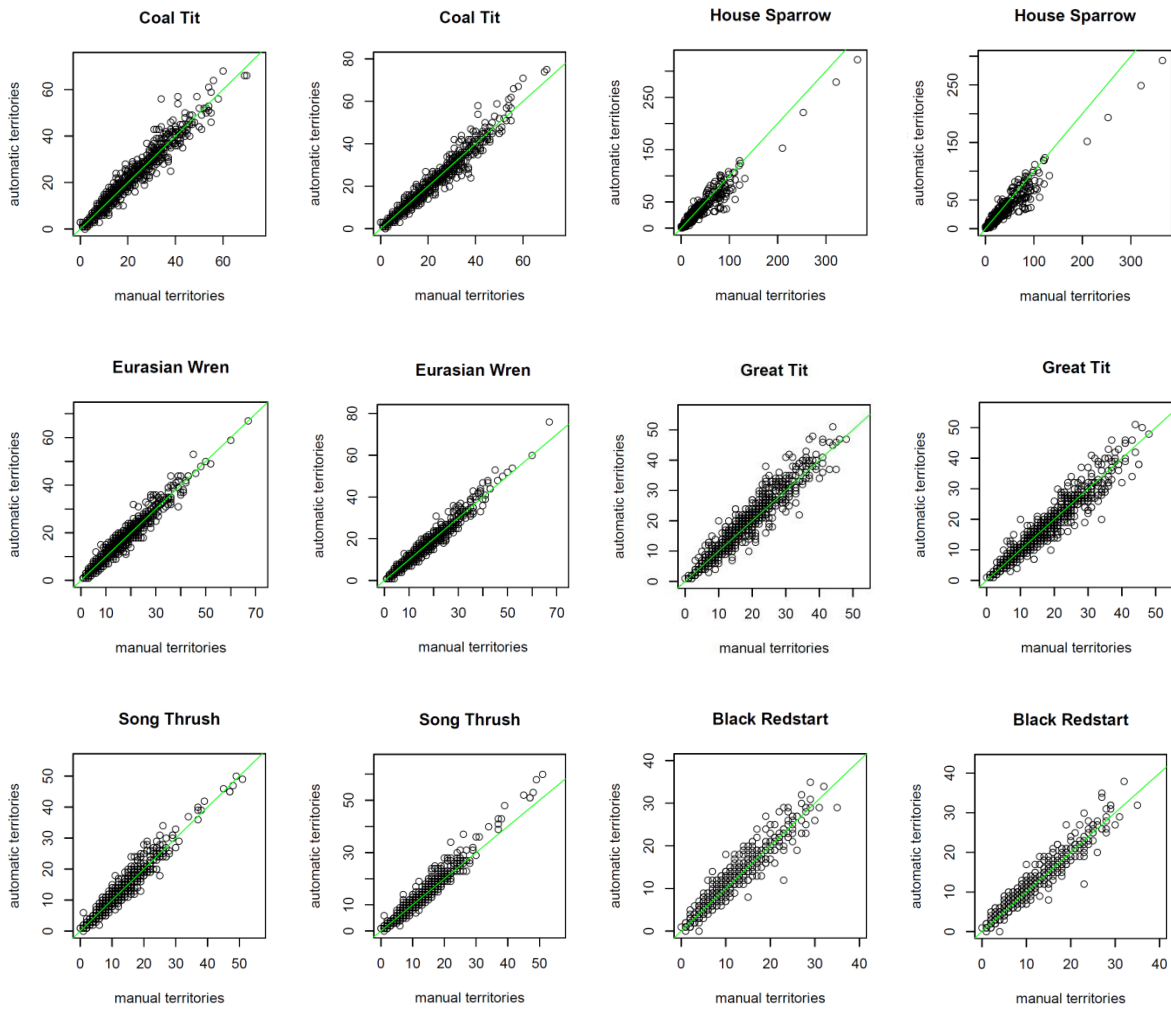
Figure 18: Number of manually delimited territories vs. number of automatically delimited territories for the 10 most common species for the Quantile Value approach (left) and the Mean Nearest Neighbour (right) approach. Each point represents one sample site in a particular year. The green line is indicating the ideal case when all automatic territory delimitations are identical to the manual solution.
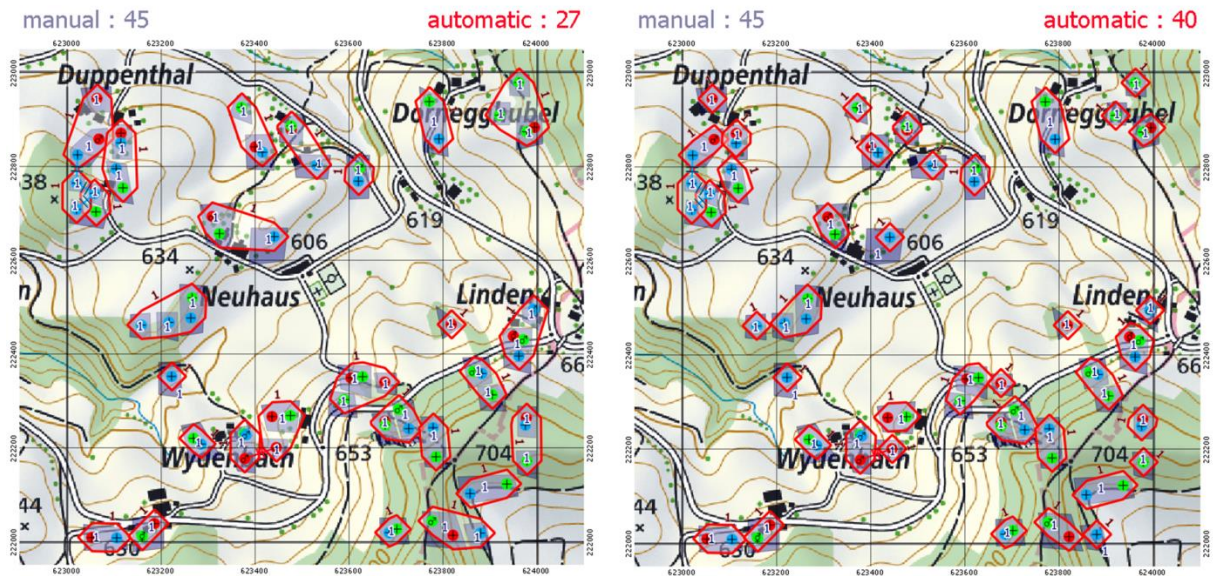
Figure 19: Automatic (red) and manual (transparent blue) territory delimitation for the Common Chaffinch in a randomly chosen sample site. The two plots show the results for the Linear Regression (left) and the Quantile Value & the Mean Nearest Neighbour approach (right). The latter two produce in this case identical clusters. Remember that the number of territories is the important end product of territory mapping (given on top of each map). Size and position of territory-polygons is only relevant for illustration but not further analyses. Base maps by swisstopo.

While previous results in this chapters only covered accuracies of the ten most common species table 11 depicts accuracy values for the ten species that worked best and worst with respect to precision (MAE_{sp}). Rare species with less than 200 manually delimited territories in 2013-2016 were excluded. Best MAE_{sp}-performance was measured for the Garden Warbler *Sylvia borin* in the Mean Nearest Neighbour approach. The Eurasian Blackcap was the most common species under the top ten. On the other hand the Common Swift *Apus apus* showed worst precision and the Common Starling was the most common species out of the worst ten.

Table 11: Results for species that perform best (upper half) and worst (lower half) with respect to their minimal precision in either Quantile Value or Mean Nearest Neighbour approach. The species set was limited to species where territory mapping data is actually used to calculate national species trends and to species that had at least 200 manually delimited territories in all the sites across all years analysed. Scientific names for species not previously mentioned: Spotted Flycatcher *Muscicapa striata,* Western Bonelli's Warbler *Phylloscopus bonelli,* Red-backed Shrike *Lanius collurio,* Dunnock *Prunella modularis,* Lesser Whitethroat *Sylvia curruca*, Eurasian Siskin *Carduelis spinus*, Common House Martin *Delichon urbicum*, Barn Swallow *Hirundo rustica*, Common Buzzard *Buteo buteo*, Red Kite *Milvus milvus*, Mallard *Anas platyrhynchos*, Carrion Crow *Corvus corone corone*.

Species	Count manual	Linear Regression MAE_{sp}	Quantile Value MAE_{sp}	Mean Nearest Neighbour MAE_{sp}	Linear Regression ME_{sp}	Quantile Value ME_{sp}	Mean Nearest Neighbour ME_{sp}
Garden Warbler	1305	0.089	0.068	0.042	0.027	-0.038	-0.002
Spotted Flycatcher	1322	0.077	0.082	0.061	0.014	-0.048	0.006
Western Bonelli's Warbler	1891	0.120	0.072	0.069	0.085	0.009	0.033
Goldcrest	6647	0.118	0.075	0.070	-0.023	0.008	-0.016
Eurasian Blackcap	20734	0.153	0.078	0.071	-0.054	0.019	-0.017
Common Redstart	762	0.113	0.077	0.072	0.092	-0.028	0.055
Red-backed Shrike	402	0.149	0.082	0.073	0.119	-0.014	0.044
Dunnock	6926	0.130	0.084	0.074	0.048	0.001	0.008
Eurasian Wren	14378	0.155	0.084	0.076	0.035	0.019	-0.003

Lesser Whitethroat	825	0.112	0.076	0.083	0.073	-0.011	0.049
Eurasian Siskin	304	0.259	0.196	0.198	0.137	0.052	0.055
Common Starling	3897	0.238	0.220	0.201	0.150	0.050	0.030
Common House Martin	2157	0.245	0.203	0.217	0.132	0.036	0.057
Barn Swallow	2472	0.251	0.208	0.212	0.172	0.063	0.079
Common Buzzard	987	0.293	0.217	0.229	0.243	0.157	0.170
Red Kite	510	0.293	0.224	0.247	0.264	0.193	0.217
Red Crossbill	787	0.331	0.252	0.283	0.260	0.117	0.165
Mallard	714	0.287	0.254	0.272	0.177	0.092	0.136
Carrion Crow	3693	0.367	0.264	0.278	0.353	0.226	0.240
Common Swift	1138	0.322	0.295	0.304	0.153	0.093	0.108

Table 12 summarises the findings across all species: The Quantile Value approach revealed the best results for most species. The results of the Mean Nearest Neighbour approach were however not far away. For the Quantile Value approach 86 species showed a MAE_{sp} of <15%. These species make up 76.6% of all territories. For the Mean Nearest Neighbour approach 80 species had a MAE_{sp} below 15% which corresponds to 77.4% of all territories. While 77 species showed a bias of less than 5% (ME_{sp}) for the Quantile Value approach, for the Mean Nearest Neighbour approach this held true for 59 species. The Linear Regression approach was always performing substantially worse.

Table 12: Summarised performance of the three approaches. Values indicate the number of species for which an approach was performing best or below a specific accuracy value (see chapter 4.1.3 for a discussion of acceptable error). Analysis was run for 162 species. If two approaches performed equally well for a particular species, the species was counted for both approaches.

	Number of species for which an approach performed best		
	Linear Regression	Quantile Value	Mean Nearest Neighbour
Number of species with lowest MAE_{sp}	49	105	99
Number of species with lowest $abs(ME_{sp})$	56	128	69
Number of species with MAE_{sp}<15%	55	86	80
Number of species with $abs(ME_{sp})$<5%	42	77	59

3.1.4 Comparison with manual territory delimitation
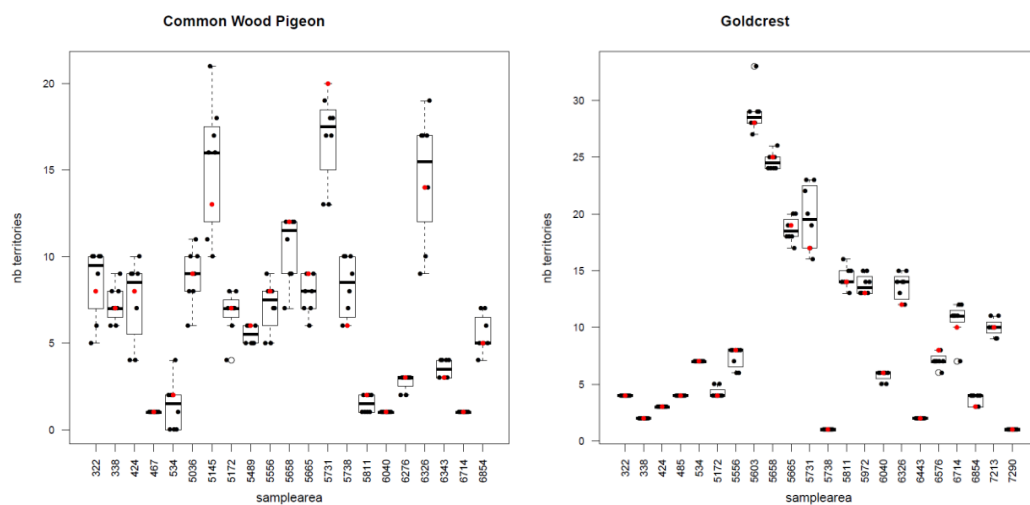
Figure 20: Number of delimited territories by seven different experts (boxplots and black dots) and the original user (red dots) for different test sample sites. Plots for all seven species are found in appendix 7.4.

The so far presented accuracies are difficult to interpret without knowing about the error that is inherent to manual territory delimitation. We know that there is considerable variation when delimitation is done by different people (see figure 2 or figure 20 for illustration). In this chapter we'll compare inaccuracies due to manual analysis (analyst effect) with inaccuracies found in automatic analysis. Figure 21 and figure 22 give a first impression on how the number of manually delimited territories compared to the number of automatically delimited territories. They show how well automatic solutions (black) fitted to the original user's solution across all analysed sample sites. Plotted on top (red) is the comparison of the experts' territory delimitation to the original user. The plots show for both the automatic and experts' analyses comparable scatter along the green line that would represent perfect match between original user and automatic or expert analysis.
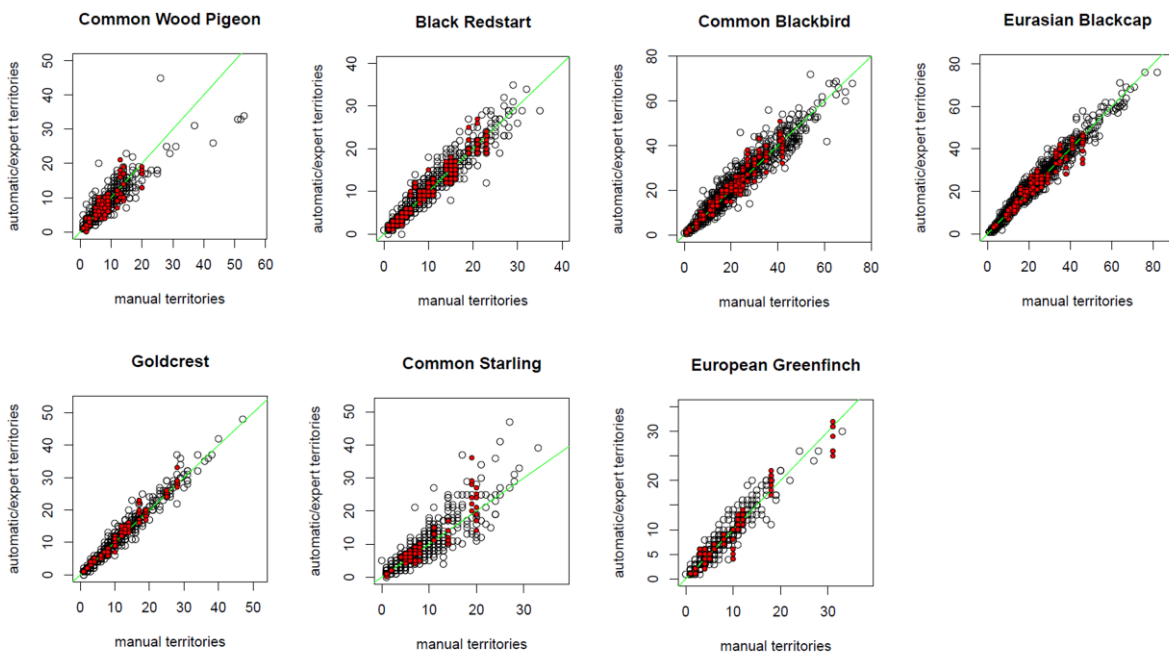
Figure 21: Direct comparison of results from the Quantile Value approach (black circles) and manual territory delimitation by experts (red dots). The x-axis shows the number of territories delimited by the original user. The y-axis shows the number of territories delimited by the algorithm or the experts respectively. Each point corresponds to one sample site. The green line indicates perfect match between the original user's solution and the algorithm/experts.
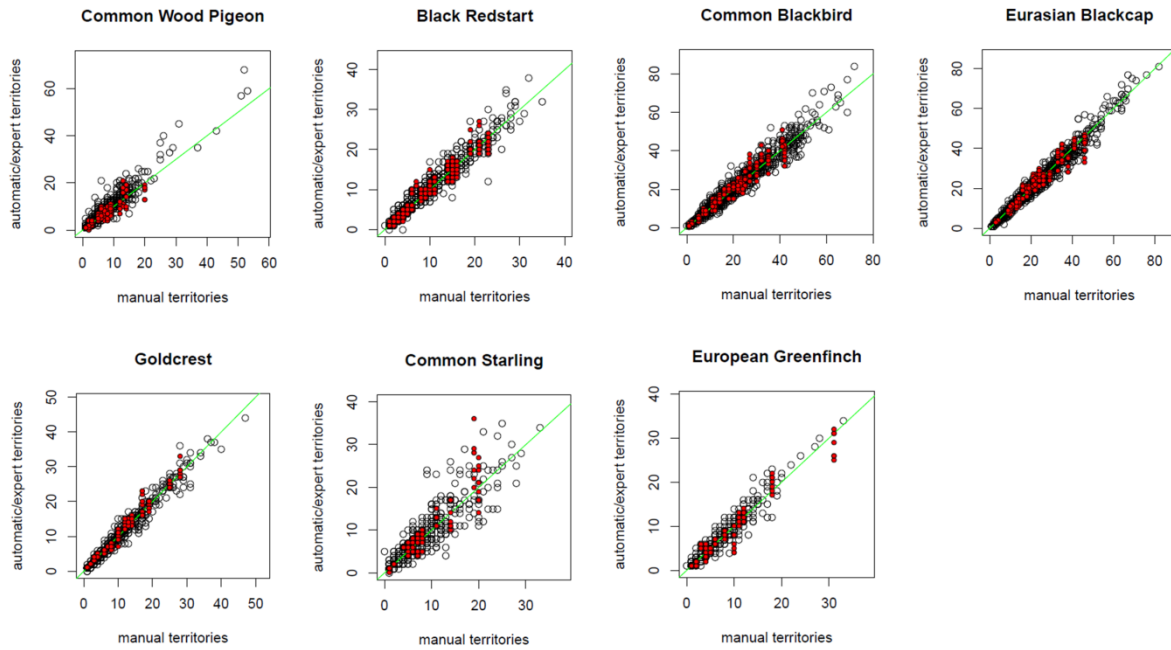
Figure 22: Direct comparison of results from the Mean Nearest Neighbour approach (black circles) and manual territory delimitation by experts (red dots). The x-axis shows the number of territories delimited by the original user. The y-axis shows the number of territories delimited by the algorithm or the experts respectively. Each point corresponds to one sample site. The green line indicates perfect match between the original user's solution and the algorithm/experts.

The distribution of the deviations from the original users territory count shown in figure 21 and figure 22 is visualised with histograms in figure 23. The shape of an ideal histogram would be narrow (indicating high precision) and centred at 0 (indicating no systematic bias, i.e. no over or underestimation of the number of territories). For automatic solutions always two histograms are depicted: once observed deviations across all sample sites and once deviations in the 35 test sample sites. The analysis of the automatic approaches in the 35 sample sites is based on a limited sample size because the seven species considered hardly ever occurred in all the 35 test sample sites. In fact only the Common Redstart was present in all 35 sites, while the Common Starling at the other end of the spectrum was only present in 14 sample sites. Figure 23 shows that histogram shapes of automatic solutions (white) are always close to or even more ideal than what is observed in expert manual territory delimitation (red) for Black Redstart, Common Blackbird, Eurasian Blackcap, Goldcrest. For the Common Wood Pigeon, the Common Starling and the European Greenfinch some of the automatic solutions show somewhat deviating distributions. Generally over-all distributions (across all sites) resemble the experts' accuracy distributions more closely than result in the 35 test sample sites for these three species.

Figure 23: Distribution of ME for seven species when analysed automatically or manually by experts. The top two figures show results for all sample sites or the 35 test sample site when analysed automatically by the Quantile Value approach. Plot 3 and 4 from top show the same for automatic analysis with the Mean Nearest Neighbour approach. The bottom plots of each page show (in red) the accuracies observed in manual analysis by experts. Values depicted in the top right corner of each plot are Mean absolute error (MAE_{sp}) and Mean Error (ME_{sp}).

European Greenfinch

Quantile Value (all sites)　　MAE = 0.114
　　　　　　　　　　　　　　ME = 0.011

Quantile Value (35 test sites)　MAE = 0.111
　　　　　　　　　　　　　　ME = −0.111

Mean Nearest Neighbour (all sites)　MAE = 0.109
　　　　　　　　　　　　　　ME = 0.021

Mean Nearest Neighbour (35 test sites)　MAE = 0.121
　　　　　　　　　　　　　　ME = 0.017

Experts (35 test sites)　MAE = 0.129
　　　　　　　　　　　　　ME = −0.02
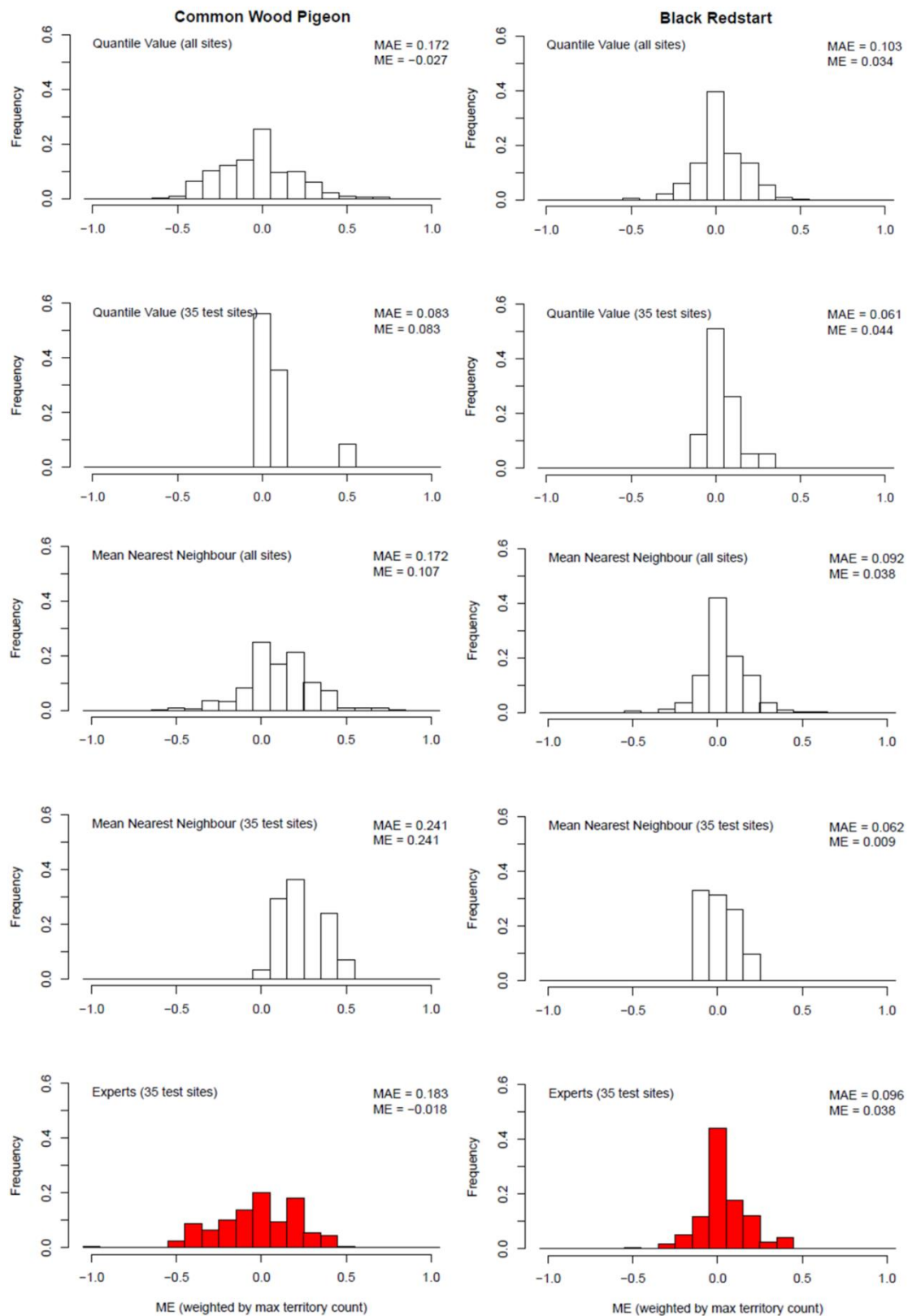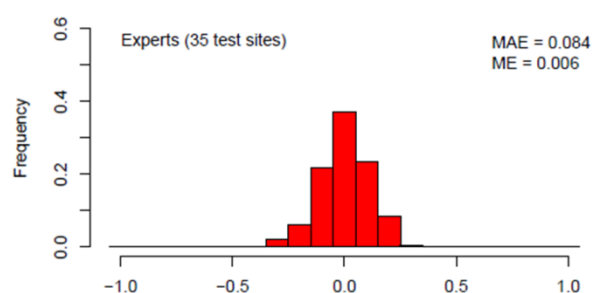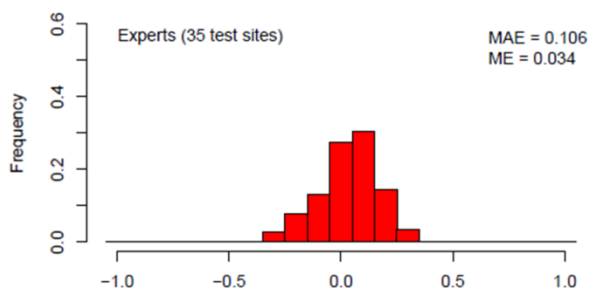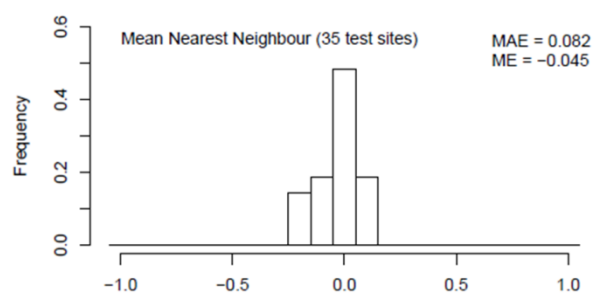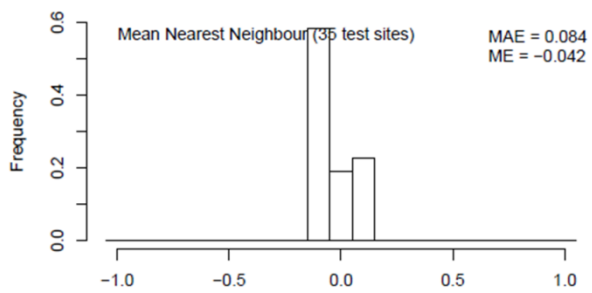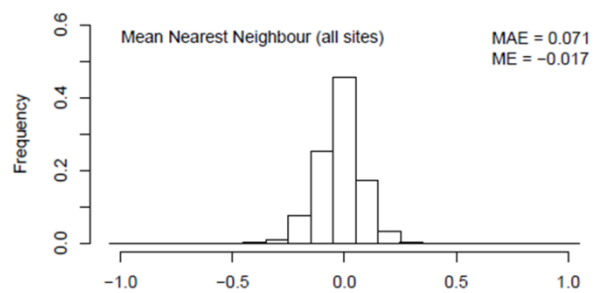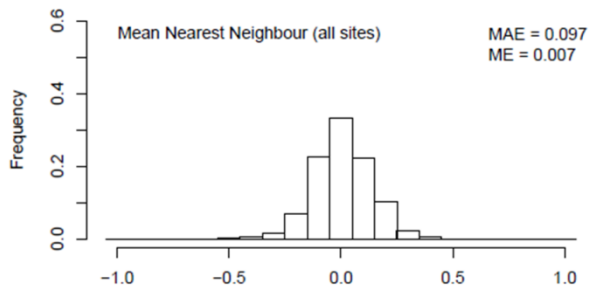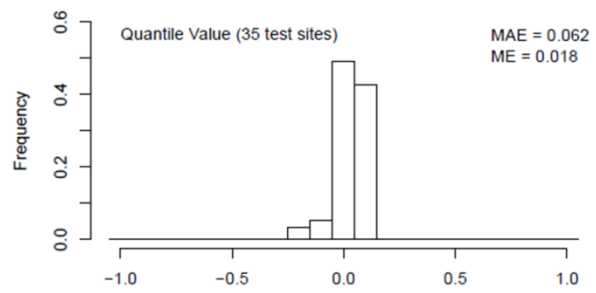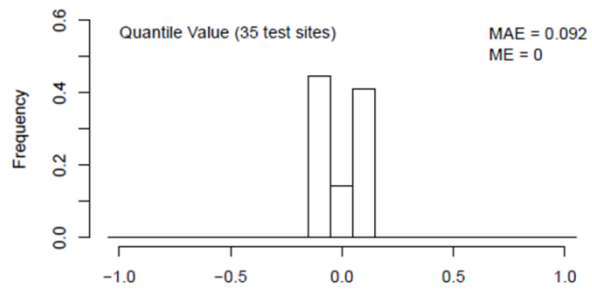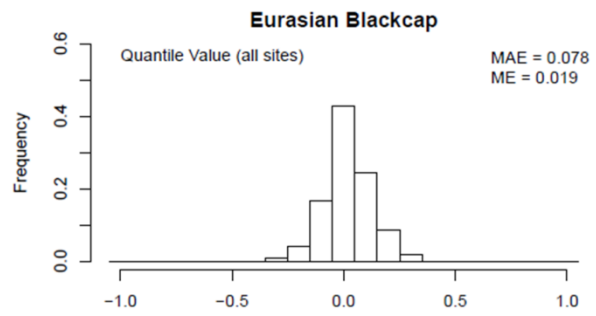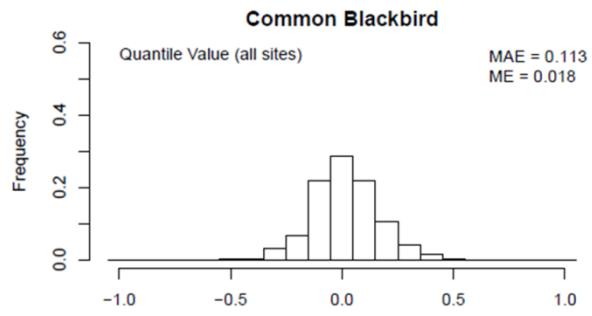
ME (weighted by max territory count)

49
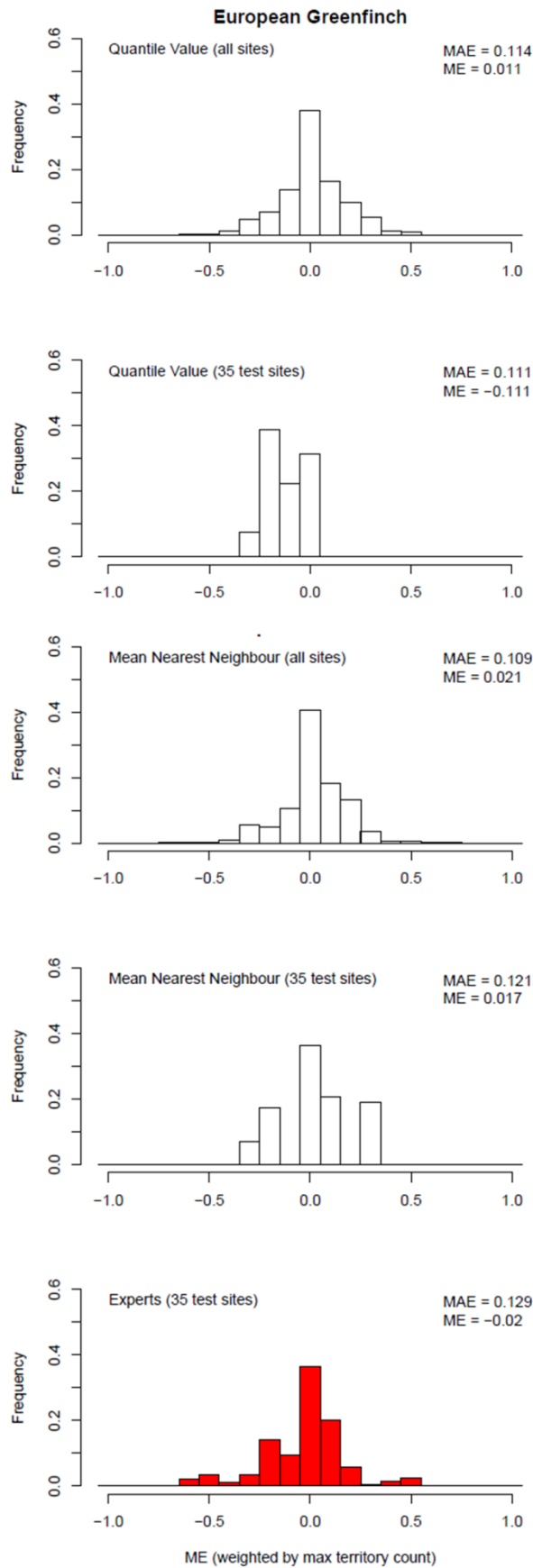
Table 13 summarizes the accuracies shown above when automatic and expert territory delimitation are applied on the same 35 sample sites (exact same input data). Table 14 compares the expert accuracy found in the 35 test sample sites with the accuracy of automatic solution across all sample sites. The precision values (MAE_{sp}) of automatic solutions were often quite similar to those found in manual territory delimitation and ranged between 0% and 25.0% depending on the species. For the species expected to be easier to analyse (Black Redstart, Common Blackbird, Eurasian Blackcap, Goldcrest) accuracy values were very close to the original user's solution and regularly even closer than observed in the experts' solution. Bias values (ME_{sp}) were (by definition) considerably smaller but deviated in some automatic analyses more heavily from 0 than the experts solutions where deviation was never larger than 8% even for difficult species. The worst automatic solution revealed a bias of 25% (Quantile value approach within the 35 test sample sites). Biases observed for some species under some automatic solution were less extreme when analysing all sample sites. Extreme values were obviously compensated to some extent in other sample sites.

Table 13: Accuracies in 35 test sample sites comparing expert territory delimitation with automatic solutions. Comparison was always conducted against the original user's solution. Values in bold are smaller or equal to what was found in the experts solutions.

Species	Expert			Quantile Value (35 test sites)		Mean Nearest Neighbour (35 test sites)	
	CV_{sp}	MAE_{sp}	ME_{sp}	MAE_{sp}	ME_{sp}	MAE_{sp}	ME_{sp}
Common Wood Pigeon	0.203	0.183	-0.018	**0.083**	0.083	0.241	0.241
Black Redstart	0.103	0.096	0.038	**0.061**	0.044	**0.062**	**0.009**
Common Blackbird	0.105	0.106	0.034	**0.092**	**0.000**	**0.084**	-0.042
Eurasian Blackcap	0.082	0.084	0.006	**0.062**	0.018	**0.082**	-0.045
Goldcrest	0.072	0.070	0.028	**0.000**	**0.000**	0.029	-0.029
Common Starling	0.175	0.191	0.080	0.250	0.194	0.197	**0.030**
European Greenfinch	0.126	0.129	-0.020	**0.111**	-0.111	**0.121**	**0.017**

Table 14: Accuracies in all automatically analysed sample sites comparing expert territory delimitation with automatic solutions. Comparison was always conducted against the original user's solution. Values in bold are smaller or equal to what was found in the experts solution.

Species	Expert			Quantile Value (all sites)		Mean Nearest Neighbour (all sites)	
	CV_{sp}	MAE_{sp}	ME_{sp}	MAE_{sp}	ME_{sp}	MAE_{sp}	ME_{sp}
Common Wood Pigeon	0.203	0.183	-0.018	**0.172**	-0.027	**0.172**	0.107
Black Redstart	0.103	0.096	0.038	0.103	**0.034**	**0.092**	**0.038**
Common Blackbird	0.105	0.106	0.034	0.113	**0.018**	**0.097**	**0.007**
Eurasian Blackcap	0.082	0.084	0.006	**0.078**	0.019	**0.071**	-0.017
Goldcrest	0.072	0.070	0.028	0.075	**0.008**	**0.070**	**-0.016**
Common Starling	0.175	0.191	0.080	0.220	**0.050**	0.201	**0.030**
European Greenfinch	0.126	0.129	-0.020	**0.114**	**0.011**	**0.109**	0.021

3.2 Computation times

The average computation time (table 15) for one sample site was between 12 (Quantile Value approach) and 32 seconds (Mean Nearest Neighbour approach). The analysis was run in R-Studio 1.1.453 (RStudio Team, 2016) using the R-Software 3.5.0 (R Core Team, 2018). The code was run on a personal computer with 16 Gigabyte RAM and an Intel Core i5-3570 CPU with 4 cores at 3.4 GHz. It was accessing a PostgreSQL 9.5 database running on a Linux Server (Ubuntu 16.04).

Table 15: Average computation times for the analysis of one sample site.

Approach	Average duration [s]
Linear Regression	13
Quantile Value	12
Mean Nearest Neighbour	32

3.3 Sensitivity to habitat and route length

This chapter evaluated whether automatic analyses varied in relationship to habitat covariates and route length. This could become relevant if covariates changed during the course of long term monitoring schemes that typically run for many decades. Five covariates and the route length calculated for each 1x1 km sample site were analysed: Percentage of area covered in forest ("forest"), area of all buildings in square meters ("buildings"), percentage of area covered in farmland ("farm"), length of all roads in meters ("roads"), elevation in meters above sea level ("elevation") and length of the route (i.e. the transect) covered during field work in meters ("route length"). For details on the calculation refer the Swiss breeding bird atlas (Knaus et al., 2018).

The bias (ME_s) of the Quantile Value approach did not seem to be influenced by forest, buildings, farm, roads nor elevation. The 95% confidence interval of the slope of the linear regression always included 0. There was however a weak negative relationship with the length of the route. If route length increased from the smallest to the longest length of routes the global bias for that sample site would "increase" by -5.6% (figure 24). The bias of the Mean Nearest Neighbour approach seemed more influenced by these covariates (figure 25). There was a weak negative relationship with forest, buildings and route length. Strongest relationships were found with roads length (slope m = -0.105) and the elevation (m = 0.118).
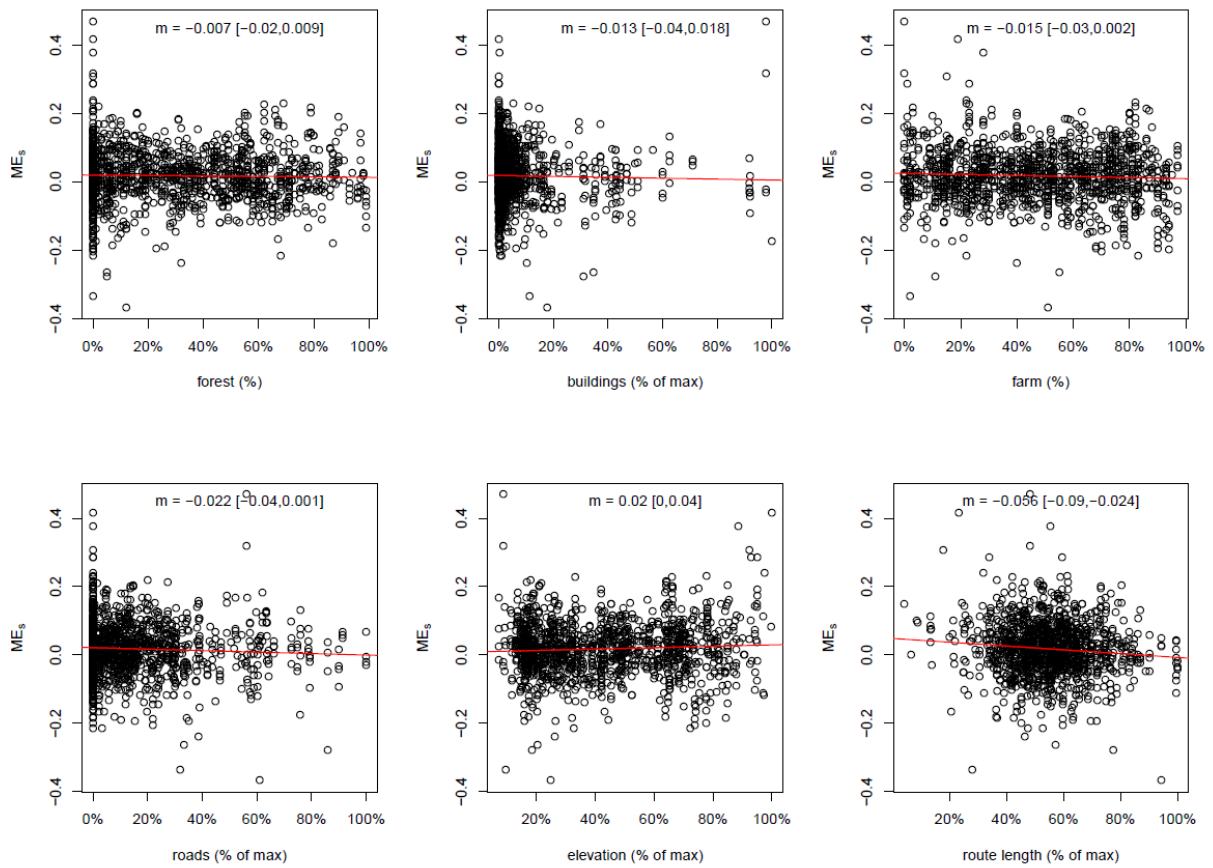
Figure 24: Relationship between habitat/route length and bias per sample site (ME_s) for the Quantile Value approach. Each point represents one sample site. The red line shows the linear regression of ME_s on the respective covariate. The slope of the linear regression and its 95% confidence interval (in brackets) are given on top of each graph. Forest and farm represent the share of the sample site covered in these habitat types. Buildings (m²), roads (m), elevation (m) and route length (m) were standardised by their maximum value.
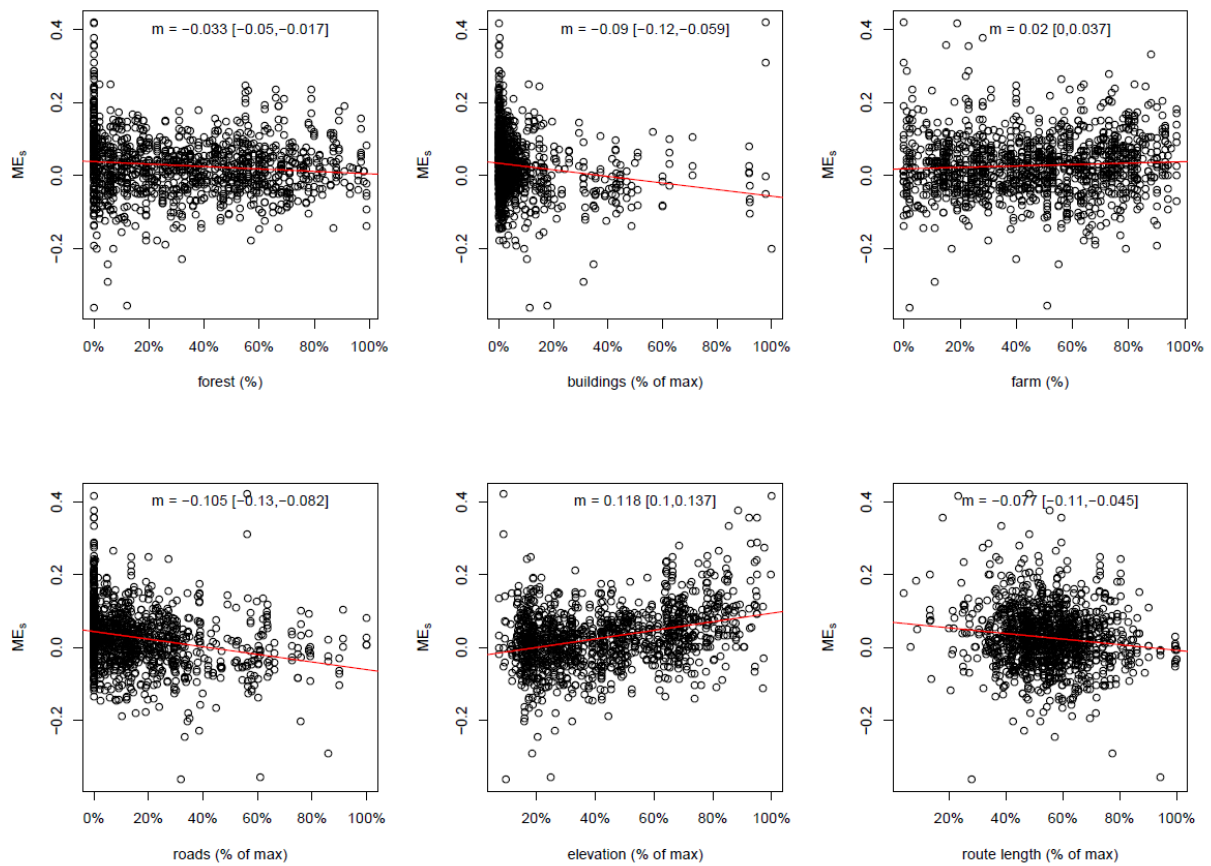
Figure 25: Relationship between habitat/route length and bias per sample site (ME_s) for the Mean Nearest Neighbour approach. Each point represents one sample site. The red line shows the linear regression of ME_s on the respective covariate. The slope of the linear regression and its 95% confidence interval (in brackets) are given on top of each graph. Forest and farm represent the share of the sample site covered in these habitat types. Buildings (m²), roads (m), elevation (m) and route length (m) were standardised by their maximum value.

Precision (MAE_s) of the automatic solutions seemed only weakly influenced by habitat or route length. Precision did not systematically change with the length of the route or the elevation for both approaches. There were however rather weak relationships between precision and forest, buildings, farm and roads for both the Quantile Value (figure 26) and the Mean Nearest Neighbour approach (figure 27).
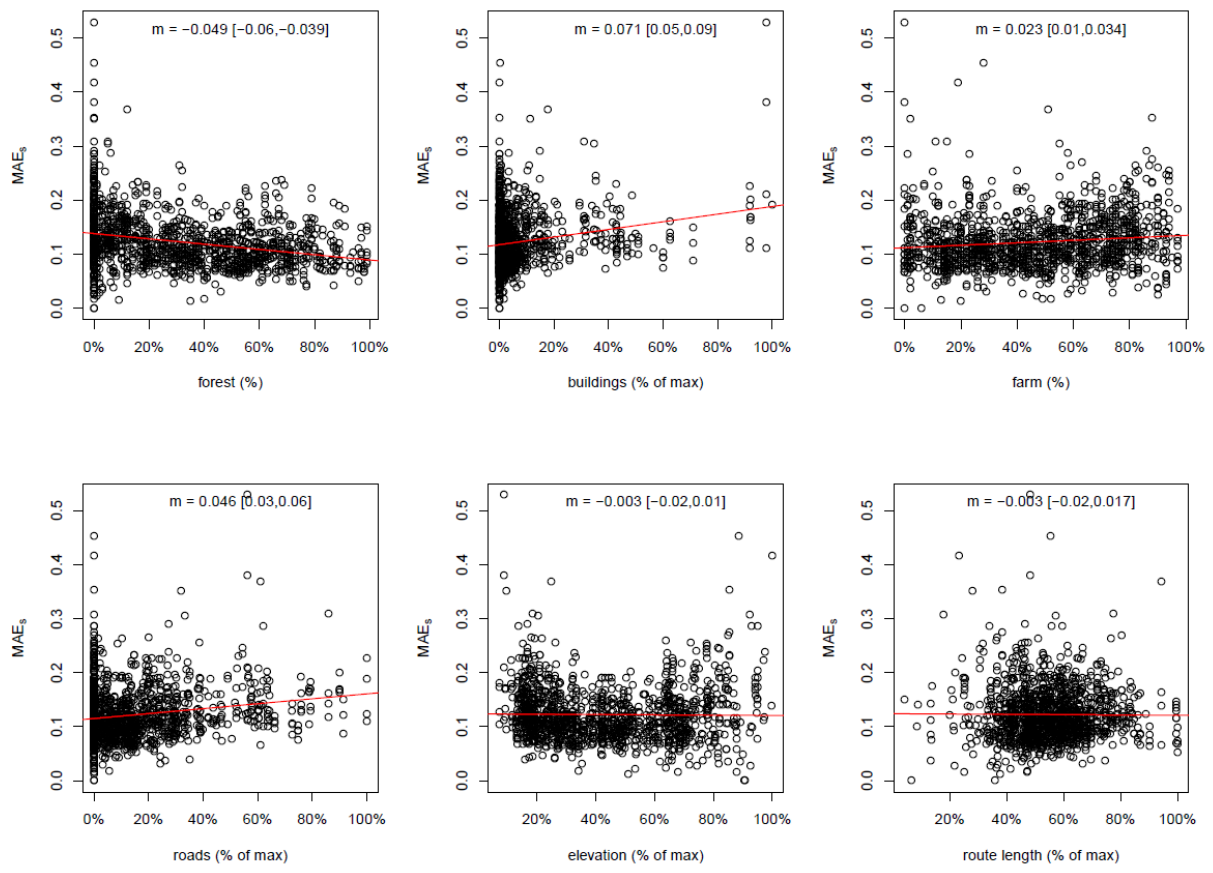
Figure 26: Relationship between habitat covariates and precision per sample site (MAE_s) for the Quantile Value approach. Each point represents one sample site. The red line shows the linear regression of MAE_s on the respective covariate. The slope of the linear regression and its 95% confidence interval (in brackets) are given on top of each graph. Forest and farm represent the share of the sample site covered in these habitat types. Buildings (m²), roads (m), elevation (m) and route length (m) were standardised by their maximum value.
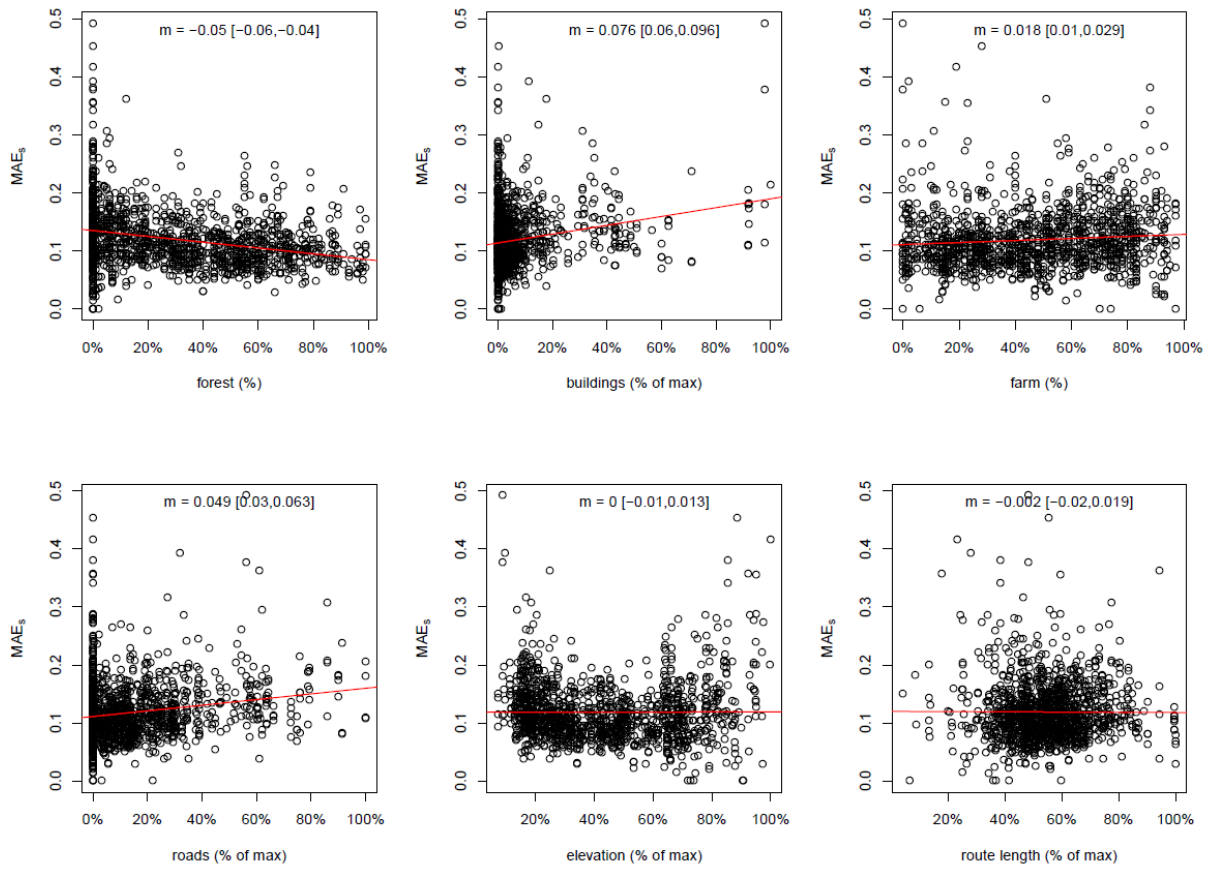
Figure 27: Relationship between habitat covariates and precision per sample site (MAE_s) for the Mean Nearest Neighbour approach. Each point represents one sample site. The red line shows the linear regression of MAE_s on the respective covariate. The slope of the linear regression and its 95% confidence interval (in brackets) are given on top of each graph. Forest and farm represent the share of the sample site covered in these habitat types. Buildings (m²), roads (m), elevation (m) and route length (m) were standardised by their maximum value.

3.5 Sensitivity to parameters

During several steps Autoterri relies on different parameters. These parameters were mentioned and explained in detail in the methods section and listed again in table 16 to give an overview about all parameters. The sensitivity of automated territory analysis to four potentially influential parameters was analysed in the following chapters.

Table 16: Overview about parameters and values used in the two most promising approaches (Mean Nearest Neighbour, Quantile Value) of Autoterri.

Name	Current value(s)	Species-specific	Potential influence	Remark
quantile	0.01-0.8	yes	very strong	Only used in Quantile Value approach to choose the normalisation distance
min_distance	20-300 m	yes	strong	Used to restrict possible normalisation distances
max_distance	100-1000 m	yes	strong	Used to restrict possible normalisation distances; used as normalisation distance when no other information available
Factor for sexually incompatible observations	1.5	no	strong	Used to increase Euclidian distances between observations that are incompatible
Factor for simultaneous observations	100	no	none	Used to increase Euclidian distances between simultaneous observations
Factor for flight observations	0.67	no	weak	Used to reduce Euclidian distances of observations to flight observation
Factor for potential double observations	0.8	no	weak	Used to decrease Euclidian distances for potential double observations
Factor for nests	100	no	none	Used to increase Euclidian distances between nests
Factor for double observations	0	no	none	Sets Euclidian distances between double observations to 0

3.5.1 Sensitivity to sexual incompatibility factor

The factor increasing the distances between sexually incompatible observations is a potentially influential parameter in the algorithm because it is applied to many distances. In order to understand better how strongly this factor influences the outcome the analysis across all sites was repeated with four different factor values ranging from 1.25 to 3. The Mean Nearest Neighbour approach was applied for this analysis.

Across all species the precision changed from at best 11.9% to at worst 12.2% deviation from the original user's solution (table 17). While the precision improved with increasing factor values the opposite was observed for overall bias. Bias was at lowest 0.4% with a factor of 1.25 and at most 4% with a factor of 3. Increasing factor values reduced the amount of underestimation more strongly then it increased the amount of overestimation (figure 28): While the median value remained more or less stable the lower whisker and the lower end of the box (1st quartile) changed more strongly into the direction of the median than the upper equivalents (whisker and 75% quantile) moved away from it. This resulted in increased precision at the price of higher general overestimation (bias).

Table 17: Response of accuracy to different factors for modifying the distances between sexually incompatible observations. Given are accuracy values resulting from the analysis with the Mean Nearest Neighbour approach.

Accuracy measure	Increase of sexually incompatible distances			
	1.25x	1.5x	2x	3x
MAE_{algo}	0.122	0.120	0.119	0.119
ME_{algo}	0.004	0.023	0.032	0.040
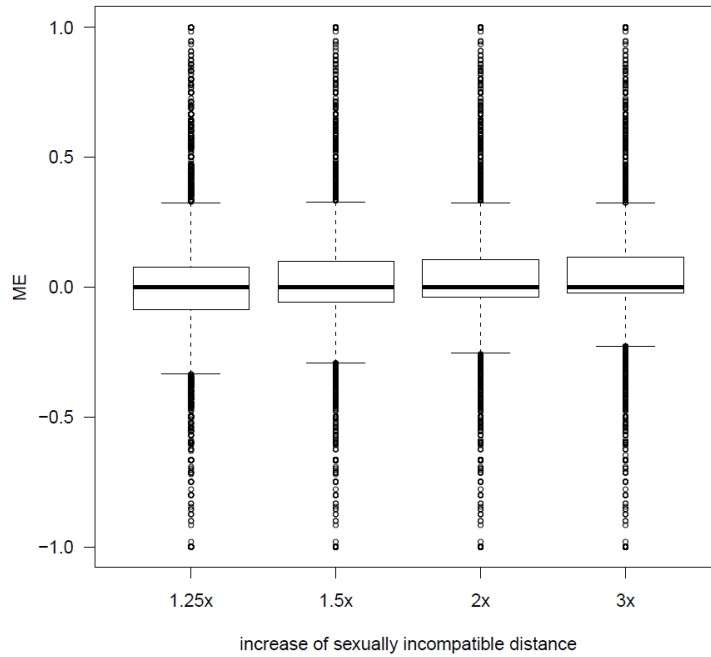
Figure 28: Response of mean error to changing factors modifying the distances between sexually incompatible observations. Boxplots show the distribution of mean errors (ME) for all species and sample sites. The four boxplots summarizes results for four different factor values. The value implemented in the algorithm presented in this thesis was 1.5.

The changes observed for the entire algorithm were also common when looking at individual species. Generally the above described patterns were regularly found in common species (e.g. Coal Tit, Common Chiffchaff, Goldcrest or Dunnock). There were however also species that didn't follow this pattern. The Common Wood Pigeon for example showed worse bias and precision (wider box) with higher factor values (see figure 29).
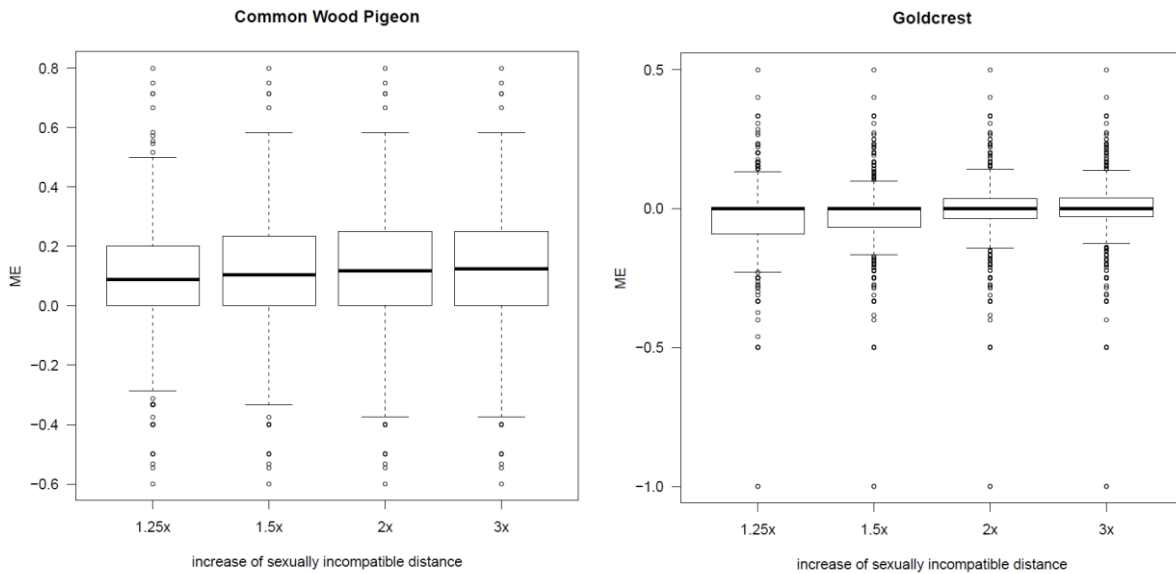
Figure 29: Response of mean error (*ME*) to changing factors modifying the distances between sexually incompatible observations for two species – Common Wood Pigeon and Goldcrest. The four boxplots summarise results for four different factor values. The value implemented in the algorithm presented was 1.5.

The modest influence of this factor is also exemplarily illustrated by direct comparisons with the manual solution (figure 30). While for the Common Wood Pigeon there was no change evident only few extreme values were influenced by the parameter in the Goldcrest.
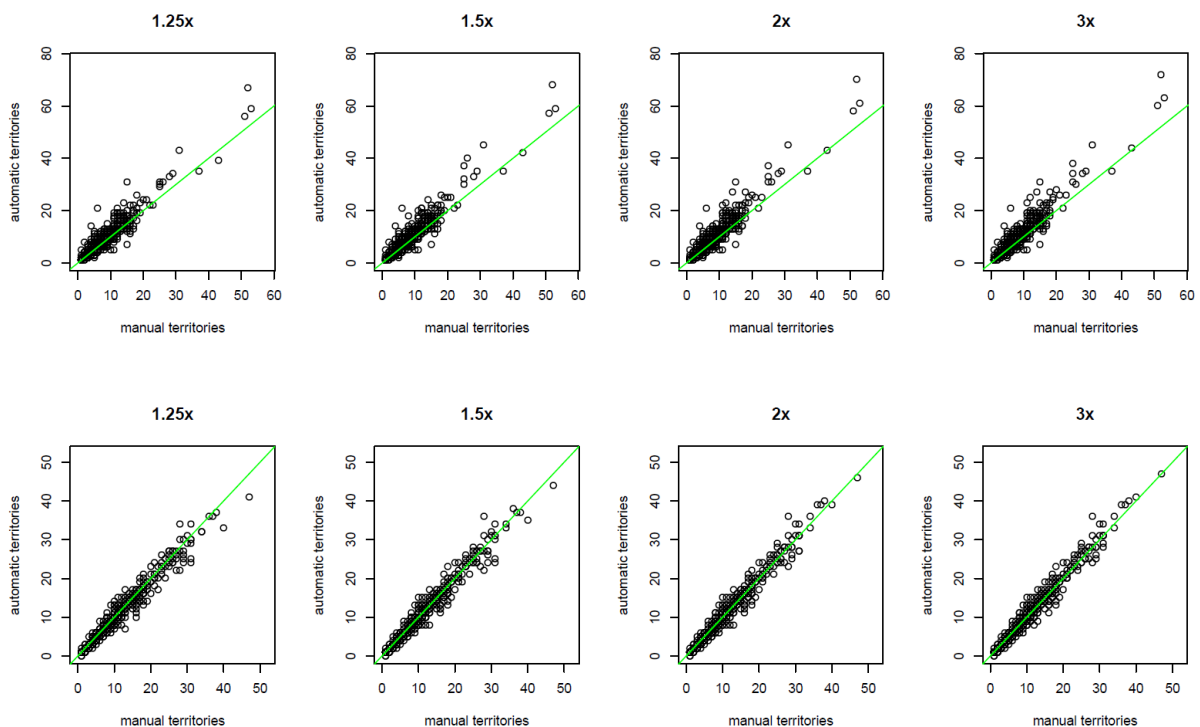
Figure 30: Comparison of automatic to manual solution with changing factors for modifying the distances between sexually incompatible observations for two species – Common Wood Pigeon (upper series) and Goldcrest (lower series).

3.5.2 Sensitivity to minimum and maximum distance

The maximum and minimum distances are used as default normalisation distance when data-based calculation of it fails. In the Mean Nearest Neighbour approach it is additionally used to limit distances that go into the calculation of the mean of the nearest neighbour distances. The influence of these distances was evaluated for one species that is common and showed a rather high bias (in order to estimate the potential to improve results) and for one common species already performing well: The Common Wood Pigeon and the Common Chaffinch.

Table 18: Accuracy of Autoterri (Mean Nearest Neighbour approach) for the Common Wood Pigeon and the Common Chaffinch under different minimum and maximum distance values.

Accuracy	min_distance Common Wood Pigeon				max_distance Common Chaffinch			
	40 m	50 m	100 m	150 m	100 m	150 m	200 m	250 m
MAE_{sp}	0.176	0.172	0.148	0.153	0.140	0.085	0.076	0.084
ME_{sp}	0.111	0.107	0.050	-0.019	0.132	0.051	0.008	-0.022

Both examples showed clearly that the species-specific distances are very influential. The overall bias (ME_{sp}) in the Common Wood Pigeon was reduced from 11.1% to -1.9%. This corresponds to an overall reduction of the number of delimited territories from 5402 to 4648. The precision (MAE_{sp}) changed less strongly and varied between 17.6% and 14.8% (table 18). By increasing the maximum distance the number of Common Chaffinch territories was continuously reduced: Best performance – ME_{sp} and MAE_{sp} closest to 0 – was achieved for a maximum distance of 200 m. In both species modifications of the distances influenced the bias more strongly than the precision: While positions of medians in figure 31 changed visibly the box width was not altered dramatically.
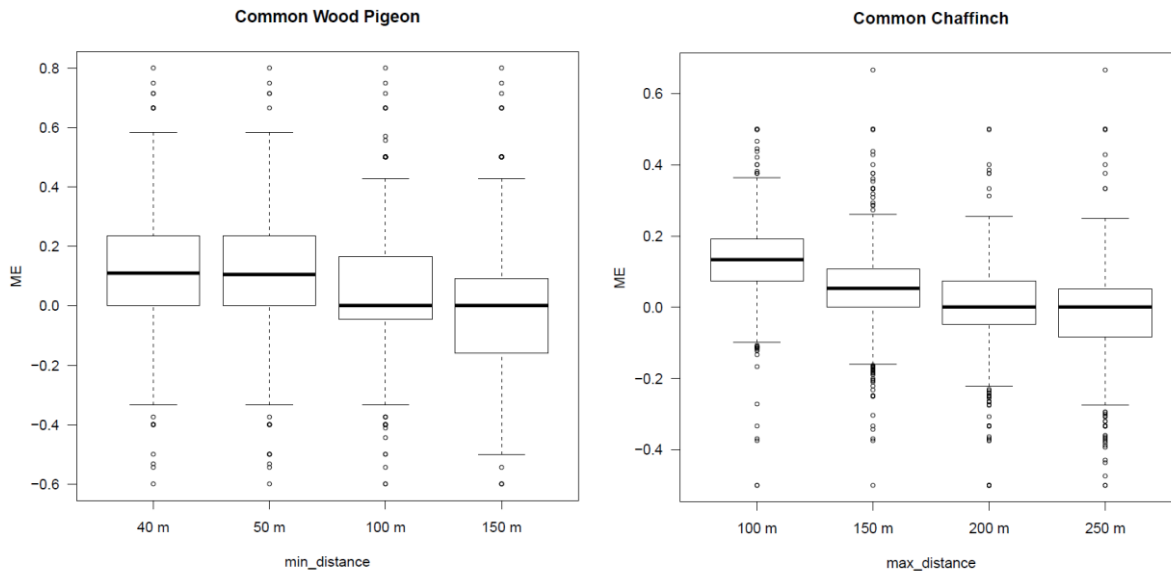
Figure 31: Sensitivity of accuracy to changing distance values for the Common Wood Pigeon and the Common Chaffinch. Boxplots show the distribution of mean errors (ME) across all sample sites for four different minimum and maximum distance values respectively. The value implemented in the algorithm presented in this thesis was 50 m (Common Wood Pigeon) and 200 m (Common Chaffinch). For this analysis the Mean Nearest Neighbour approach was applied.

Figure 32 shows the same data depicted as in the boxplots above. Interestingly the raise of the minimum distance seemed to influence sample sites with high territory counts (i.e. with dense and small territories)

the most - visible in the plots of the Common Wood Pigeon (upper four plots in figure 32). On the other hand increasing max_distance was more influencing sample sites with median territory densities while the sample sites with high densities were not affected a lot. This is not surprising as high density plots will have a lot of very short nearest neighbour distances and the influence of the filtering (to below max_distance) becomes negligible.
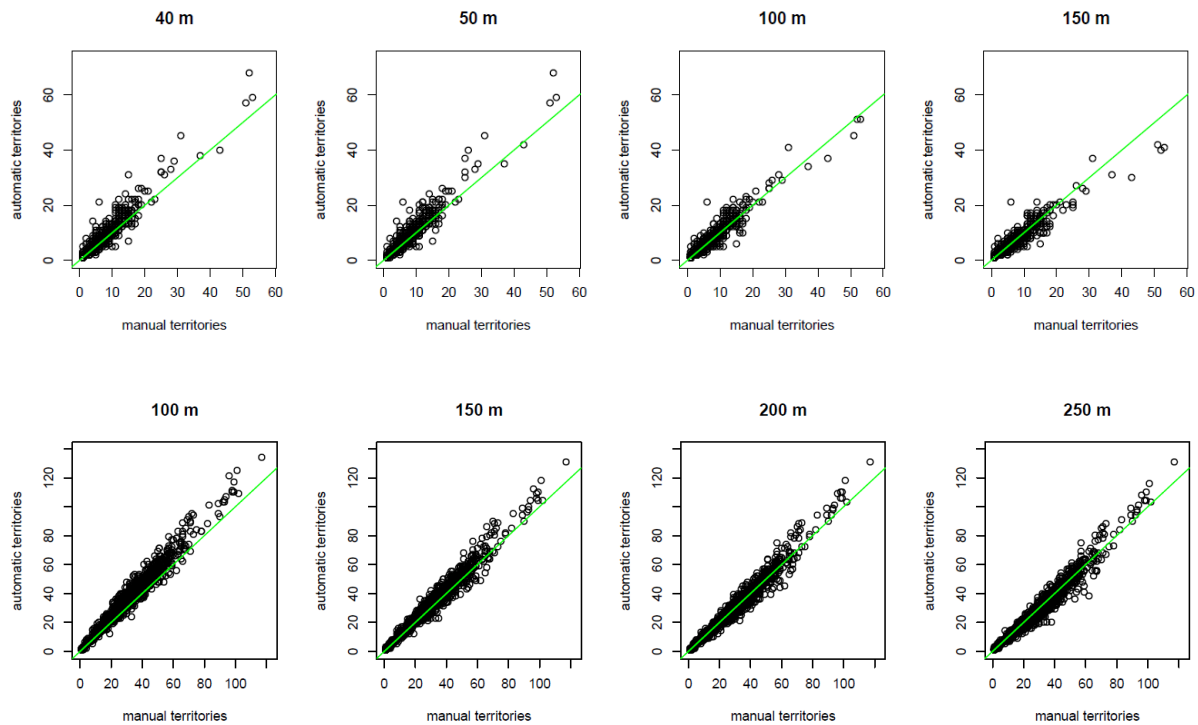
Figure 32: Sensitivity of automatic solutions to changing minimum distance values for the Common Wood Pigeon (upper series) and changing maximum distance for the Common Chaffinch (lower series). The scatterplots show how the comparison of automatic vs. manual analysis changed with different minimum/maximum distance values used in the automatic analysis. For this analysis the Mean Nearest Neighbour approach was applied.

3.5.3   Optimizing the quantile value for the Quantile Value approach

The species-specific quantile used in the Quantile Value approach is directly influencing the normalisation distance. Without any doubt the quantile value has in important influence on the performance of the Quantile Value approach. This chapter evaluates how much accuracy was improved by choosing optimal quantile values.

Across all species precision improved little from 12.9% error to 12.1%. The bias was halved from 6% to 3% after optimization (table 19).

Table 19: Improvement of mean absolute error and mean error for the entire algorithm due to optimized quantiles in sample sites of 2014 and 2016

before optimization		after optimization		improvement in %	
MAE_{algo}	ME_{algo}	MAE_{algo}	ME_{algo}	MAE_{algo}	ME_{algo}
0.129	0.060	0.121	0.030	6.2%	50.4%

While overall improvement was modest for some species the improvement was quite substantial (table 20). For the Common Chaffinch for example precision improved from 11% error before optimization to 8.4% after optimization. The bias lowered from 9.5% to 0.7%. Figure 33 shows two examples how error developed with different quantile values – for many species there was a clear optimum identifiable.

Table 20: Improvement of accuracy for the 10 most common species due to optimized quantiles in sample sites of 2014 and 2016.

species	before optimization		after optimization		improvement in %	
	MAE_{sp}	ME_{sp}	MAE_{sp}	ME_{sp}	MAE_{sp}	ME_{sp}
Common Chaffinch	0.110	0.095	0.084	0.007	24	93
Eurasian Blackcap	0.098	0.079	0.078	0.032	21	60
Common Blackbird	0.122	0.076	0.105	0.026	14	66
European Robin	0.087	0.026	0.087	0.026	0	0
Coal Tit	0.115	0.083	0.091	0.017	21	80
Eurasian Wren	0.086	0.037	0.083	0.020	4	45
House Sparrow	0.174	-0.061	0.174	-0.063	0	-3
Great Tit	0.124	0.083	0.111	0.061	10	26
Song Thrush	0.098	0.034	0.096	0.023	3	33
Black Redstart	0.100	0.046	0.100	0.046	0	0
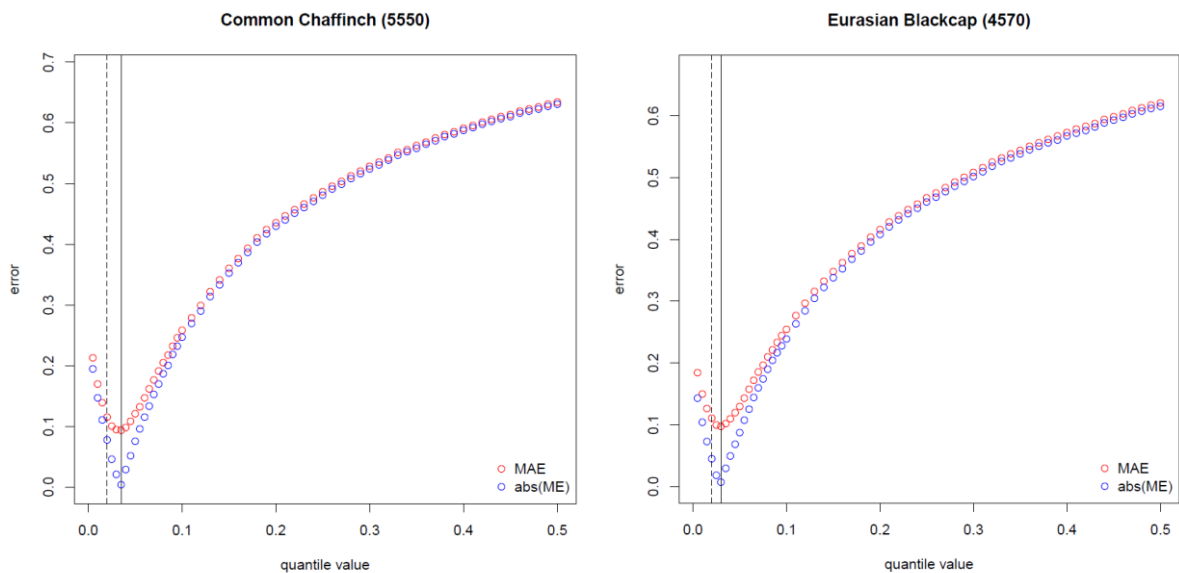
Figure 33: Values for the mean absolute error (red) and the mean error (blue) as a function of the quantile value for the two most common species. The data-derived optimum (solid line) and manually chosen quantile (dashed line) were close but not identical illustrating the potential for improvement.
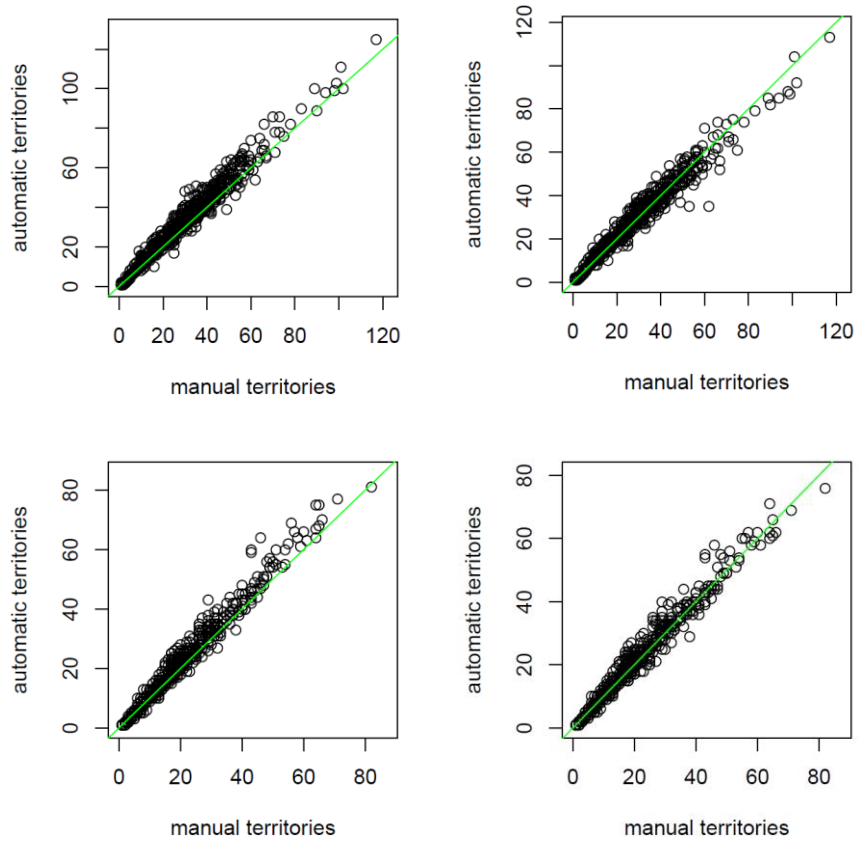
Figure 34: Comparison of territory counts for the years 2014 and 2016 from manual and automatic territory delimitation before (left) and after (right) quantile optimization. Figures show the situation for Common Chaffinch (top), Eurasian Blackcap (bottom).

4 Discussion

4.1 Accuracy of Autoterri

4.1.1   Substantial variation between "easy" and "difficult" species

The automatic delimitation of territories revealed striking performance differences between species (see for example table 11 or appendix 0). While for some species it was easily possible to reproduce similar territory counts as in manual territory delimitation for other species this was difficult. For the 10 best performing species (Garden Warbler, Spotted Flycatcher, Western Bonelli's Warbler, Goldcrest, Eurasian Blackcap, Common Redstart, Red-backed Shrike, Dunnock, Eurasian Wren, Lesser Whitethroat) precision (MAE_{sp}) of the Quantile Value und Mean Nearest Neighbour approach was always below 8.5%. For the 10 worst-performing yet still common species (Eurasian Siskin, Common Starling, Common House Martin, Barn Swallow, Common Buzzard, Red Kite, Red Crossbill, Mallard, Carrion Crow, Common Swift) precision ranged between 19.6% and 30.4%. In some exceptionally difficult species like the Alpine Chough *Pyrrhocorax graculus* precision was as bad as 44.6% (appendix 0).

These differences in precision between species are however little surprising as they reflect a common pattern: Well-performing species are all highly territorial and more or less monogamous during one breeding season. This is not the case for species performing badly. They usually don't defend a territory or at most only a small area around their nest site. Species like the Common House Martin, Barn Swallow, Common Swift and to some extent also Common Starling and Eurasian Siskin may all form colonies or lose aggregations where many pairs can breed in very close vicinity. And very obviously all of these species fly around a lot (for example for visiting temporarily available food sources like a freshly mown meadow). These characteristics make consistent territory delimitation at least difficult if not impossible because they violate a basic assumption of the territory mapping method: territoriality of birds. The warning in Bibby et al. (2000) and Voříšek et al (2008) that territory mapping is inefficient for non-territorial species, semi-colonial species, those that range wide, or those that are not monogamous is aiming exactly at this violation of the basic assumption and can be confirmed with the findings presented above.

This problem is however not limited to automatic analysis but is equally prevalent in manual territory delimitation (see also discussion in O'Connor and Marchant, 1981). A finding the comparison of automatic vs. manual territory delimitation in this thesis confirmed: Common Starling, European Greenfinch and Common Wood pigeon, three species specifically chosen for the analysis because they stress the assumption of territory mapping, showed worse precision values than clearly territorial species in automatic solution but also in manual territory delimitation by experts (see chapter 3.1.4 and discussion in the following chapter 4.1.2).

4.1.2   Comparison with manual territory delimitation

Comparing manual territory delimitation to the two most promising approaches of Autoterri (Quantile Value and Mean Nearest Neighbour) quickly revealed that accuracy values are often surprisingly similar - a first but strong indication that either of these two approaches have the potential to replace manual territory delimitation. Compared to manual territory delimitation by experts automatic solutions were often closer to the results of the original user: The Quantile Value approach performed better than the experts with respect to precision in six out of the seven species investigated (table 13). The bias was

however only closer to 0 in two of the seven species. The Mean Nearest Neighbour approach was more precise in five and less biased in three species. All in all values differed only by a few percent compared to manual delimitation by experts. It seems that inaccuracies measured are not a characteristic of the algorithms but are to a large extent inherent uncertainty in the data that cannot be interpreted more accurately even by a powerful machine like the human brain.

There were two distinct exceptions in the comparison of the 35 test sites: The Mean Nearest Neighbour approach overestimated the number of territories for the Wood Pigeon considerably compared to the analysis by experts (ME_{sp} of 24.1% compared to -1.8%). And the Quantile Value approach showed a substantially higher overestimation for the Common Starling (ME_{sp} of 19.4% compared to 8%). While the species-specific quantile in the Quantile Value Approach has been subject to optimisation (see chapter 0 and 3.5.3) the Mean Nearest Neighbour approach was never subject to an optimization process. All parameters and values were chosen based on the author's opinion. While it will be more difficult to optimize the Mean Nearest Neighbour approach as it has not one very influential value like the quantile, there is room for optimization. Modifications in the species-specific minimum and maximum distance values are promising starting points (also for further improvement of the Quantile Value approach). In fact, for the Common Wood Pigeon it has for example been shown in chapter 3.5.2 that a modification of the minimum distance value from 50 m to 150 m reduced the global bias of the Mean Nearest Neighbour approach from 10.7% to -1.9%. This example is illustrating the possibility and need for optimisation of parameters.

The behaviour of the automatic territory delimitation was analysed for seven species – once for the 35 test sample sites to allow for a non-confounded comparison with the experts solution and once for all 1268 surveys in the data set. It is striking how strong accuracy values sometimes changed when comparing the analysis of 35 sites with the complete data set (compare table 13 and table 14): While overall error for the Mean Nearest Neighbour approach for the Common Wood Pigeon in the 35 sites was for example 24.1% it was only 10.7% across all sample sites (i.e. the 1268 surveys). The differences could probably be explained by the small sample size. Usually not all 35 sample sites analysed were occupied by a species. The Common Wood Pigeon for example only occupied 21 of the 35 test sample sites. The accuracy calculation for automatic solution for the 35 test sample sites was hence based on only 21 data points. The delimitation by the experts was repeated 7 times per sample site so the data size was bigger, i.e. 147 data points for the Common Wood Pigeon. While simple chance effects may play some role the differences nevertheless show that there may be substantial variation between sites. This is also visible in the histograms of

figure 23 – distributions are rather wide for those species that show big changes between the analysis across 35 test sample sites vs. the overall data set.

Because the true number of bird territories is unknown automatic delimitation aims at copying manual territory delimitation as closely as possible. But because manual delimitation allows for different interpretations it is impossible to reach the exact same result. So inevitably there will be differences but how much difference is acceptable? The comparison of the manual vs. automatic territory delimitation is really the only option for evaluating how well an automatic approach is behaving. This became evident for the Common Starling where MAE_{sp} of the Mean Nearest Neighbour approach was 20.1%: a value that seems at first sight unacceptably high. However even the experts showed substantial deviation from the original users solution for the Common Starling. Their MAE_{sp} across 35 test sample sites was 19.1% and hence comparably high. So while we should potentially be prepared to accept bad precision values as long as they are comparable with manual solutions, there are probably some limits to what seems acceptable for bias values (ME_{sp}). Something discussed in the following chapter.

4.1.3   What are acceptable accuracy values?

The analysis of automatic solutions provides us with accuracy and bias values. But only for seven species I was able to directly compare with inaccuracies measured in manual territory delimitation. For the other species the question remains if the found precisions and biases are acceptable. How much variation is normal in manual analysis and hence also acceptable for an automatic territory analysis?

How much territory numbers vary when different people analyse the same observations was quantified in several studies. The variation was usually expressed with the coefficient of variation (CV). A measure strongly correlated to the MAE used in this thesis (see 2.3.3.3). Scheffer (1987) found across 16 species maps analysed by 14 ornithologist a CV of 21.6%. Svensson (1974) calculated a CV of 21.3% when 58 persons analysed 37 species maps of 6 species. Another trial in the same study revealed a CV of 20% when 17 persons analysed six species maps. Best (1975) obtained a CV of 23% for one species analysed by 5 different people. Verner and Milne (1990) found in two different sites CV of 14.6% and 27.2%. On the other hand O'Connor and Marchant (1981) found in an analysis of the British Common Birds Census only for two out of 26 species significant analysts effects. The median CV over all species was only 5% in their study with very well trained and experienced analysts. They also showed that CVs vary from species to species and also depending on the person that conducted the field work. CV for the Common Linnet *Carduelis cannabina* for example varied from 13.3%-34.4% depending on the field worker. CV for the Chaffinch only varied between 2.4% and 3.7%. In a field project on disturbance of Yves Bötsch (personal communication, July 2018) repeated analysis by three different people revealed CV of 20.3% for Great Tits and 13.3% for Blue Tits.

All these publications reveal some general patterns: First of all, across all these studies average CV values are astonishingly similar and vary around 20% (with the exception of the findings of O'Connor and Marchant (1981)). The average values found across the seven species analysed in the here presented study was 12.4% (if weighted by the number of territories CV=10.7%). The fact that CVs in O'Connor and Marchant (1981) and this study are lower than in other studies is probably a consequence of rigorous checking and validation implemented in the common breeding bird monitoring schemes in Britain and Switzerland. It is also not surprising that values in O'Connor and Marchant (1981) are yet slightly lower than in Switzerland as the British mapping scheme is based on 8-10 visits while in Switzerland it's only 2-3 visits. With more visits it is expected to have more precise information about the number of territories. Secondly CVs are highly species-specific. For one thing this makes comparison between different studies

difficult. It was also shown that species that behave highly territorial, don't fly around a lot and don't breed in colonies are easiest to analyse consistently (see 4.1.1 and also discussion in O'Connor and Marchant, 1981).

Of 99 species that rely at least partially on territory mapping data for the calculation of their Swiss national population trend 30 or 34 (Quantile Value or Mean Nearest Neighbour approach respectively) show MAE_{sp}-values below 0.1, 58 or 55 show values below 0.15, 79 or 78 below 0.2 and 95 or 94 below 0.3. In comparison to error values found in literature and in the manual territory delimitation these values seem by no means extreme and indicate that a large part of the species could actually be analysed automatically with respect to overall precision. Four species however never reached MAE_{sp} values below 0.3 neither for the Quantile Value nor the Mean Nearest Neighbour approach: Corn Bunting *Emberiza calandra*, Meadow Pipit *Anthus pratensis*, Common Grasshopper Warbler *Locustella naevia* and Willow Warbler *Phylloscopus trochilus*. Interestingly the last three species were also mentioned in van Dijk et al. (2013) as problematic species because manual territory delimitation was not applied consistently (Gerard Troost, personal communication, August 2018). These four species are well-known cases where it is notoriously difficult to distinguish migrants/vagrants from local breeding birds. The decision whether one saw a migrant or a local breeding bird leaves substantial scope for subjective decisions in manual territory delimitation, which is probably the reason for the unprecise automatic territory delimitation.

To summarize the previous reflections I argue, based on direct comparison with manual territory delimitation and literature values, that acceptable species-specific precision (MAE_{sp}) values could be as high as 15% for stereotypically territorial species and up to 30% for species that stress the territoriality assumption of the method. With respect to overall bias (ME_{sp}) it is probably somewhat different. The maximal bias (ME_{sp}) in manual territory delimitation by experts was 8% for the Common Starling. For all other species values were below 4%. Even if it's difficult to arrive at the same number of territories as the original user we'd expect that an algorithm is not badly over- or underestimating the total number of territories across all sample sites. This could induce biases in subsequent analyses (e.g. jumps in species population trends). For species like the Red Kite that show an overestimation of 19% there is a urgent need for further optimisation. During development of the algorithm presented the focus was primary on the abundant species. While it was for example possible to improve ME_{sp} values for the ten most common species on average by 40% by finding an optimum quantile in the Quantile Value approach (see chapter 3.5.3), this optimisation process was ineffective for species like the Red Kite that show little simultaneous observations and occur in low densities (i.e. have little between territory distance information in the data). In these cases all approaches of Autoterri regularly make use of the species-specific minimum and maximum distance (just as a human analyst would). By modifying these distances there is room to reduce biases seen in many of these species and improve overall results of Autoterri substantially.

4.2 Computation time

Average computation time per sample site varied between slightly more than 10 sec (Linear Regression and the Quantile Value approach) and 32 sec (Mean Nearest Neighbour). As these are average values, it probably takes more than 1 min to analyse extremely bird-rich sample sites with the Mean Nearest Neighbour approach (exact computation times for each site were not monitored).

If Autoterri were to be used productively it would either be implemented as a button "delimit territories now" in the user interface of the Web GIS used to analyse the data or run automatically after volunteers

have digitized their field data completely (trigger Autoterri upon status change). Waiting for 1 min would probably be acceptable for most volunteers considering the analysis time of approximately 1 hour that they save. More drastic is the potential gain for the Swiss Ornithological Institute that is currently investing roughly 50 man-days for validating manual territory delimitations. If manual delimitation was completely replaced by automatic delimitation checking could be reduced dramatically. Alternatively automatic solutions could be used to identify situations where manual territory delimitation should be checked and hence help improve checking efficiency.

Nevertheless some thought should be put into speeding up the algorithm once it is decided to actually use it productively. So far no effort was invested in optimising computation time. Possible options to check would be the following: First, optimize R script where possible. Secondly, evaluate if procedural language PL/R (Conway, 2009) that allows to write PostgreSQL function and triggers in the R programming language is more performant. Thirdly, see how performance is improved if Autoterri is run on servers instead of a personal computer. Fourthly, consider rewriting the entire code in a more performant programming language. Options one to three are probably inevitable anyway when implementing the algorithm.

4.3 Robustness of Autoterri

4.3.1 Sensitivity to habitat and route length

Precision and bias of automatic solutions are obviously correlated to some habitat variables and often also route length. This is potentially relevant for a long term monitoring program if these variables change. Long term bird monitoring wants to monitor changes in bird population sizes but not changes in habitat. If bird populations are underestimated in cities compared to rural landscapes urban scrawl could lead to a "measured" decline in bird population that doesn't correspond to the real population changes.

Of the analysed variables not all seem problematic. While elevation does not change enough in human times scales and the length of the route can be controlled by project managers other variables (forest, buildings, farm and roads) will however change continuously. But how realistic are changes that actually exhibit an influence on the monitoring data? The strongest effect was found for the influence of roads on the bias of the Mean Nearest Neighbour approach (figure 25). If a sample site changed from a square with 0 m of roads to almost 27 km of roads (i.e. most extreme sample site in Zurich down town) the number of territories would be estimated 10% lower. While 10% is a not an irrelevant value such extreme changes in habitat are not realistic. Even an increase of road length from 0% to 10% highest value would be extreme and would only result in a 1% lower territory estimate. All in all the potential to influence results is for the current algorithms smaller than the uncertainty inherent to manual territory delimitation where bias values of 0.6%-8% (table 13) were observed. We should nevertheless keep an eye on this issue and update this analysis after modifying/optimizing parameter because it has a broad effect (i.e. on all sample sites).

In the above paragraph I only discussed the influence of covariates on bias but not on precision. While I already illustrated why an effect on bias can be problematic it seems less problematic if precision is reduced in different habitats as these changes will be much smaller than imprecision observed due to different analysis.

The Quantile Value approach seems less susceptible than the Mean Nearest Neighbour approach to being biased "by" covariates. While the Quantile Value approach is only influenced by the (controllable) route

length the Mean Nearest Neighbour results may be biased by forest, buildings, roads, elevation and route length. This may be explained by the optimization process the Quantile Value approach has experienced (leaving us with optimal quantiles) but has not yet been done for the Mean Nearest Neighbour approach. In chapter 3.5.2 it was exemplarily shown that optimizing parameters like the minimum distance can influence results of sample sites with high Common Wood Pigeon densities – high densities are very likely observed in similar habitat types. So if parameters are optimized also correlations to covariates could vanish or at least be weakened.

Finally the observed correlations between covariates and bias (ME) could either be introduced directly by the habitat (or route length) by influencing the way field workers record and analyse birds. It is for example easier to detect, separate and precisely locate birds in open spaces – something that could influence the number of territories we delimit. On the other hand the observed correlations could also appear by mere differences in species composition. If for example alpine species are more overestimated this will show up as a positive slope in the elevation plot of figure 25. A visual scan of species-specific responses to these covariates did not show an evident common pattern within a covariate hence indicating species-composition effect. But a more thorough analysis is needed to understand the actual mechanism better.

4.3.2   Sensitivity to parameters of Autoterri

Different approaches of Autoterri apply fixed or species-specific parameters. The analyses in chapter 3.4 illustrated that some of these parameters are influential on the overall results (like minimum distance, maximum distance, quantile) while others are obviously of less importance despite their regular use (factor for sexual incompatible observations). Because some are influential they absolutely need to be optimized to the data available. That this can substantially improve overall performance was illustrated for the quantile value optimization (chapter 3.5.3) that halved overestimation and improved accuracy from 12.9% to 12.1%. Nevertheless here presented accuracy and bias values are expected to not worsen but improve more or less moderately depending on the species.

Minimum and maximum distances are two parameters that contain a lot of room for optimization as has been illustrated for the Common Wood Pigeon. It seems tempting to adjust these distances so Autoterri fits best to the data. However if the minimum and maximum distance are too close together it is almost equivalent as using fixed distances for delimiting the territories. This is something that should be avoided with respect to natural changes in territory densities across time. The Common Wood Pigeon for example has become more common in the recent decades (Knaus et al., 2018), which probably lead to denser breeding i.e. smaller distances between territories. The Mean Nearest Neighbour approach for example can only take such changes into consideration if a sufficiently wide range of nearest neighbour distances is accepted for the calculation of the normalisation distance. Simultaneous observations – that will be easier to record if a species starts breeding more densely – can to some extent compensate for too restrictive nearest neighbour distances as they always generate different territories, no matter how big the normalisation distance is. I conclusion one has to be aware that for the maximum and minimum distance parameters there is a risk of losing flexibility of the algorithms to adjust for future situations by fitting parameters too optimally to current data.

While other parameters seem less influential they could still be optimized. There is however several ways to do so – I'll give a few thoughts: Optimization of the factor for flight observation could for example be optimized for species that show a large proportion of flight observations. An optimized factor for this parameter is maybe more important for Red Kite than for Common Chaffinch territory delimitation. Because potential double observations are so rare they offer however no big potential to improve the

algorithm. Sexually incompatible observations from the same visit are in some countries almost treated like simultaneous observations (van Dijk et al., 2013) – which would ask for a much higher factor for sexually incompatible observation than the currently chosen 1.5. I argue that sexually incompatible observations from the same visit should not be treated equally as simultaneous observation as it is easily possible in the field to record the same individual twice – especially for species ranging widely. And finally while it is theoretically possible to make many of these parameters species-specific I recommend being parsimonious on doing so: The more complex and specific an algorithm becomes the more difficult it gets do understand the algorithm itself and even more so consequences to it if environment or field methods change (i.e. recording by smartphone app instead of paper maps).

4.3.3   Sensitivity to field method

The influence of parameters used in Autoterri was extensively discussed and partially evaluated. The way how field data are recorded was not yet looked at. There are many influencing factors that could change the way field workers record data over time. I will just name two: Regular feedback from the Swiss Ornithological Institute pushes observers to put more attention on recording as many simultaneous observations as possible. Also the introduction of a mobile app to record observations digitally instead on paper maps is under discussion and could potentially influence input data to Autoterri. On the other hand there is also different territory mapping projects that apply more visits and/or have sample sites that are of different shape and size than the 1x1 squares analysed so far. While generally I believe that Autoterri should be more or less resistant to such differences in field method this should be verified with simulated data.

4.4 Comparison with existing algorithms

Subsequently I'll compare Autoterri with five other algorithms found in literature to analyse territory mapping data (see introduction in chapter 1.5.1 for details on these algorithms).

The algorithms of North (1977) and Gerß (1984) use the single linkage clustering method which is creating chain-like clusters – something that is undesired for bird territories. Scheffer (1987), Marchand (2015) and the here presented Autoterri improved this shortcoming by using an "average closeness" (average linkage or Wards clustering method) and not the closeness of the two most extreme points. North (1977) simply used one maximum fusion distance above which clusters could not be combined. Already Scheffer (1987) remarked that a fixed maximal fusion distances as used by North (1977) is not a feasible way to analyse territory mapping data as "the optimal value of maximum distance is strongly dependent on the species analysed as well as on the habitat that has been censused". Scheffer (1987) subsequently tried to detect the optimal distance by finding the point where clustering jumped from within-territory distances to between territory distances with his own method and Cook's distances proposed by Gerß (1984). But Scheffer (1987) abandoned the approach because finding this point "inevitably fails occasionally because of misleading structure of the data". Instead he used fixed maximum distances for two groups of species (species with large territories and species with small territories). And also Marchand's (2015) algorithm – that was a precursor of the Linear Regression approach of Autoterri – failed to detect this point sufficiently well as the comparisons in chapter 3.1 showed. Finally the algorithm of van Dijk et al. (2013) sounds promising as it is operational for several years now. Due to differences in the field work I argue that their algorithm is nevertheless not applicable to Swiss conditions: While the Netherlands apply territory mapping with 8-12 visits Switzerland is using a simplified territory mapping approach with only 2-3 visits. Due to the incredible amount of data there will be almost always a

simultaneous observation (or incompatible observations from the same visit) separating two territories at least in one out of the 8-12 visits in the Netherlands. This is not the case for the 2-3 visits in Switzerland. Many territories are not separated by simultaneous observations (nor observations of the same visit). An application of the approach of van Dijk et al. (2013) would most likely lead to severe underestimation in simplified territory mapping.

While there is five algorithms only three - North (1977), Scheffers (1987) and van Dijk et al. (2013) - provided quantitative comparison of automatic with manual solutions. North (1977) presented a table with results for different maximal distance values for the Common Blackbird and Chaffinch. From the table I chose per species the distance delivering best results and calculated the MAE_{algo} which was roughly 8.5%. This value is comparable with what is found in Autoterri. However the result of Autoterri is probably better because for one thing the MAE_{algo} for the algorithm of North (1977) is most likely suffering from overfitting as it was chosen to be optimal for one specific sample site. And secondly North (1977) based his analysis also on traditional territory mapping with up to 10 visits which could be expected to result in increased precision when compared to the analysis of simplified territory mapping data. Scheffer (1987) analysed 16 species maps with his algorithm and at the same time manually by 14 experienced ornithologist. While the average deviation by the ornithologists (expressed as CV) was 21.6% the algorithms average absolute deviation of the computed numbers from the estimations made by original user was only 5.6%! This value is only half as big at what is found in Autoterri. The difference to Autoterri may be explainable by two points: Firstly, Scheffer (1987) probably based his analysis on traditional territory mapping data, hence having data from many visits. This could reduce uncertainty in the analysis because more information about neighbourhood (especially simultaneous observations) will be available with more visits. Secondly it is unfortunately unclear what species he analysed. The choice of species will be highly relevant to the output of an overall result (see discussion in chapter 4.2). Nevertheless the strong reduction in uncertainty when analysing territories automatic instead of manually is surprising/impressive. Van Dijk et al. (2013) evaluated the influence of their "Autocluster" on species population trends and found that 12 of 92 species changed significantly compared to trend based on manual delimitation - more than expected by chance. According to their analysis Autocluster in comparison to manual territory delimitation only led to larger counts in 3 species (+9% to +19%) and smaller counts in 9 species (-10% to -31%). A direct comparison with Autoterri is currently not possible because comparisons are reported in different units.

The comparison showed that Autocluster (van Dijk et al., 2013) is most likely not applicable to the Swiss data and "historic" algorithms suffer from conceptual shortcomings (single linkage, fixed maximum distance values) that were remedied in Autoterri. Despite the difficulties in comparing quantitative accuracy values the order of magnitude (between 0-30% deviation per species) is comparable to what is found in Autoterri. In conclusion I would argue that Autoterri is currently the best available option to analyse simplified territory mapping data.

4.5 Research questions revisited

1. *Is it possible to produce one algorithm for the analysis of all species of breeding birds?*
   In principle yes it is possible. Only for 4 out of 99 species that base their trend on territory mapping data it was not possible to produce precision values below 30%. For species that stress the territoriality assumption of the territory mapping method automatic analysis is however difficult (as is manual analysis).

2. *Can an automatic analysis produce results that are comparable with results produced by manual territory delimitation?*
   Yes errors observed in manual territory delimitation by experts were of the same magnitude as automatic solutions.

3. *Are there sensitivities to environmental conditions (e.g. area of settlements) or parameter values that could pose a risk for the long term use of the developed automatic solution?*
   Yes there are correlations with habitat and/or route length parameters but they are weak in their effect strength and currently pose no risk. But they need to be reanalysed after optimization of Autoterri.
   Yes some parameters (specifically the minimum and maximum distance values) are very influential and have to be carefully chosen in an optimisation process. For long-term analysis the parameter values are not a threat as long as the span between minimum and maximum distance is sufficiently wide to allow the algorithm to adapt for changes in density.

4. *What is the computation time to analyse one sample site?*
   Average computation time per sample site varied between a bit more than 10 sec (Linear Regression and the Quantile Value approach) and 32 sec (Mean Nearest Neighbour approach).

4.6 Conclusion & perspectives

It has been shown that automatic analysis of territory mapping data by Autoterri reveals similar results like manual territory analysis. Two approaches – Quantile Value and Mean Nearest Neighbour approach – showed similar accuracy values. Personally I find the Mean Nearest Neighbour approach more promising for the following two reasons:

1) Intuitively and conceptually it is more comprehensive that the average of all distances to the next neighbour (nearest neighbour distances) is a value that could be used as a distance to separate territories. By definition these distances seem destined to differentiate territories while choosing a species-specific quantile is a more artificial approach trying to optimize fit between manual and automatic territory delimitation.
2) The most influential parameter in the Quantile Value approach - the quantile - was already optimized. The Mean Nearest Neighbour approach has not undergone an optimisation process yet and still it performs similarly. By optimizing the minimum and maximum distance the Mean Nearest Neighbour approach has a greater potential to improve than the Quantile Value approach.

In order to improve and eventually use Autoterri productively in long term monitoring I propose the following next steps

- Optimise all parameters and carefully discuss results, especially minimum and maximum results with respect to future flexibility of the Autoterri to adapt to changes.
- Re-evaluate the correlation of accuracy and habitat covariates and route length. Evaluate realistic effect strengths and evaluate if they are problematic.
- Evaluate influence of changes in the field method. What happens for example if the number of simultaneous observations is reduced? What happens if more observations per territory are recorded? What happens if sample site size was different than 1x1 km? What happens if we record birds in four instead of three visits?

- Tackle single difficult species with special rules. Find for example ways to detect observations that are related to temporal vertical movements due to late snow in alpine species ("Schneeflüchter") that should not lead to territories.
- Develop a polygon representation of the territories that is non-overlapping and where each point is only contained in one polygon.

5 Acknowledgements

6 References

Aljalbout, E., Golkov, V., Siddiqui, Y., Cremers, D., 2018. Clustering with Deep Learning: Taxonomy and New Methods. ArXiv Prepr. ArXiv180107648.

Best, L.B., 1975. Interpretational errors in the" mapping method" as a census technique. The Auk 92, 452–460.

Bibby, C.J., Burgess, N.D., Hill, D.A., Mustoe, S., 2000. Bird census techniques. Elsevier.

Bubb, P., Almond, R., Kapos, V., Stanwell-Smith, D., Jenkins, M., 2010. Guidance for national biodiversity indicator development and use. UNEP-WCMC Camb. UK 120.

Conway, J.E., 2009. PL/R User's Guide-R Procedural Language.

Dean, D.J., Thakar, V., Sirdeshmukh, N., 2015. Optimal Routefinding Across Landscapes Featuring High-cost Linear Obstacles. Trans. GIS.

Eger, K., 2018. Choosing the right linkage method for hierarchical clustering [WWW Document]. URL https://stats.stackexchange.com/questions/195446/choosing-the-right-linkage-method-for-hierarchical-clustering (accessed 11.8.18).

Eisank, C., Koch, A., 2016. Clusteranalyse.

Enemar, A., Sjöstrand, B., Svensson, S., 1978. The effect of observer variability on bird census results obtained by a territory mapping technique. Ornis Scand. 31–39.

Gerß, W., 1984. Automatische Revierabgrenzung bei Siedlungsdichteuntersuchungen. J. Für Ornithol. 125, 189–199.

Gonseth, Y., Wohlgemuth, T., Sansonnens, B., Buttler, A., 2001. Die biogeographischen Regionen der Schweiz. Erläuterungen und Einteilungsstandard. Umwelt Materialien Nr. 137. Bundesamt Für Umw. Wald Landsch. Bern 48.

Gregory, R.D., van Strien, A., 2010. Wild bird indicators: using composite population trends of birds as measures of environmental health. Ornithol. Sci. 9, 3–22.

Kéry, M., Royle, J.A., 2016. Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models. Academic Press.

Kéry, M., Royle, J.A., Schmid, H., 2005. Modeling avian abundance from replicated counts using binomial mixture models. Ecol. Appl. 15, 1450–1461.

Klvanova, A., Vorisek, P., Gregory, R., Van Strien, A., Meyling, A.G., 2009. Wild birds as indicators in Europe: latest results from the Pan-European Common Bird Monitoring Scheme (PECBMS). Avocetta 33, 7–12.

Knaus, P., Antoniazza, S., Wechsler, S., Guélat, J., Kéry, M., Strebel, N., Sattler, T., 2018. Swiss Breeding Bird Atlas 2013–2016. Distribution and population trends of birds in Switzerland and Liechtenstein. Swiss Ornithological Institute, Sempach.

Knaus, P., Schmid, H., 2014. Beispiele von Reviergrössen bei vereinfachten Revierkartierungen.

Marchand, J.-R., 2015. Analyse spatiale d'observations ornithologiques (Bachelor Thesis). Haute Ecole d'Ingénierie et de gestion du canton de Vaud, Yverdone-les-Bains.

Müllner, D., 2013. Fast Hierarchical, Agglomerative Clustering Rou- tines for R and Python. Journal of Statistical Software 1–18.

North, P.M., 1977. A novel clustering method for estimating numbers of bird territories. Appl. Stat. 149–155.

O'Connor, R., Marchant, J., 1981. A field validation of some Common Birds Census techniques. British Trust for Ornithology, Thetford.

R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Robinson, R.A., Sutherland, W.J., 2002. Post-war changes in arable farming and biodiversity in Great Britain. J. Appl. Ecol. 39, 157–176.

Royle, J.A., Kéry, M., Gautier, R., Schmid, H., 2007. Hierarchical spatial models of abundance and occurrence from imperfect survey data. Ecol. Monogr. 77, 465–481.

Royle, J.A., Kery, M., Guelat, J., 2011. Spatial capture-recapture models for search-encounter data. Methods Ecol. Evol. 2, 602–611.

Royle, J.A., Nichols, J.D., Kéry, M., 2005. Modelling occurrence and abundance of species when detection is imperfect. Oikos 110, 353–359.

RStudio Team, 2016. RStudio: Integrated Development Environment for R. RStudio, Inc., Boston, MA.

Sattler, T., Knaus, P., Schmid, H., Volet, B., 2017. Zustand der Vogelwelt in der Schweiz. Bericht 2017. Schweizerische Vogelwarte, Sempach.

Scheffer, M., 1987. An automated method for extimating the number of bird territories from an observation map. Ardea 75, 231–236.

Schmid, H., 2017. TerriMap-online – ein Tool zur rationellen Auswertung von vereinfachten Revierkartierungen. Swiss Ornithological Institute.

Schmid, H., Burkhardt, M., Keller, V., Knaus, P., Volet, B., Zbinden, N., 2001. Entwicklung der Vogelwelt in der Schweiz, Avifauna Report Sempach 1, Annex.

Schmid, H., Luder, R., Naef-Daenzer, B., Graf, R., Zbinden, N., 1998. Schweizer Brutvogelatlas. Verbreit. Brutvögel Schweiz Im Fürst. Liecht. 574.

Schmid, H., Spiess, M., 2008. Brutvogelaufnahmen bei BDM-Z7 und MHB: Anleitung zur Entscheidfi ndung bei Grenzfällen und zur Revierausscheidung.

Svensson, S., 1974. Interpersonal variation in species map evaluation in bird census work with the mapping method. Acta Ornithol.

Tan, P.N., Steinbach, M., Kumar, V., 2005. Cluster Analysis: Basic Concepts and Algorithms.

van Dijk, A.J., Noback, M., Troost, G., Vergeer, J.-W., Sierdsema, H., van Turnhout, C., 2013. De introductie van Autocluster in het Broedvogel Monitoring Project. Limosa 86 94–102.

Verner, J., Milne, K.A., 1990. Analyst and observer variability in density estimates from spot mapping. Condor 313–325.

Voříšek, P., Klvaňová, A., Wotton, S., Gregory, R.D., 2008. Best Practise Guide for wild Bird Monitoring Schemes.

Ward Jr, J.H., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58, 236–244.

7 Appendix

7.1 Quantile optimization

Table 21: Overview over the quantile optimization showing whether a better quantile could be calculated and how quantiles changed from non-optimized (before) to optimized (after). Optimization was not possible for 67 mainly rare species.

Name	manual count 2013-2016	optimized	quantile before	quantile after	remark
Mute Swan	35	yes	0.05	0.31	
Mallard	714	yes	0.05	0.49	
Hazel Grouse	88	yes	0.05	0.46	
Rock Ptarmigan	234	yes	0.05	0.26	
Black Grouse	651	yes	0.1	0.43	
Common Pheasant	15	yes	0.05	0.42	
Great Crested Grebe	59	yes	0.05	0.5	
Red Kite	510	yes	0.25	0.48	
Common Buzzard	987	yes	0.25	0.37	
Common Kestrel	656	yes	0.25	0.38	
Common Moorhen	24	yes	0.05	0.07	
Eurasian Coot	266	yes	0.05	0.21	
Feral Pigeon	773	yes	0.03	0.015	
Common Wood Pigeon	4781	yes	0.08	0.16	
Eurasian Collared Dove	450	yes	0.05	0.48	
Common Swift	1138	yes	0.2	0.41	
Eurasian Wryneck	96	yes	0.05	0.48	
European Green Woodpecker	918	yes	0.05	0.49	
Great Spotted Woodpecker	3011	yes	0.05	0.25	
Lesser Spotted Woodpecker	53	yes	0.05	0.21	
Woodlark	21	yes	0.05	0.18	
Eurasian Skylark	955	yes	0.1	0.25	
Eurasian Crag Martin	287	yes	0.05	0.34	
Barn Swallow	2472	yes	0.1	0.5	
Common House Martin	2157	yes	0.05	0.5	
Tree Pipit	2881	yes	0.1	0.095	
Water Pipit	5983	yes	0.03	0.075	
Grey Wagtail	531	yes	0.05	0.36	
White Wagtail	3057	yes	0.03	0.01	
Eurasian Wren	14378	yes	0.03	0.035	
Dunnock	6926	yes	0.03	0.05	
Alpine Accentor	848	yes	0.1	0.41	
European Robin	16134	yes	0.04	0.04	
Common Nightingale	90	yes	0.05	0.05	manually to 0.05
Black Redstart	9209	yes	0.05	0.05	
Common Redstart	762	yes	0.05	0.25	
Whinchat	443	yes	0.05	0.15	
Northern Wheatear	1569	yes	0.07	0.18	
Common Rock Thrush	80	yes	0.05	0.3	
Ring Ouzel	1939	yes	0.07	0.17	
Common Blackbird	19205	yes	0.03	0.04	
Fieldfare	1355	yes	0.05	0.47	
Song Thrush	9766	yes	0.05	0.055	
Mistle Thrush	5120	yes	0.05	0.03	
Eurasian Reed Warbler	290	yes	0.05	0.075	
Lesser Whitethroat	825	yes	0.05	0.27	
Garden Warbler	1305	yes	0.05	0.08	

Eurasian Blackcap	20734	yes	0.02	0.03	
Western Bonelli's Warbler	1891	yes	0.05	0.07	
Wood Warbler	243	yes	0.05	0.47	
Common Chiffchaff	8274	yes	0.05	0.04	
Willow Warbler	113	yes	0.05	0.46	
Goldcrest	6647	yes	0.04	0.035	
Common Firecrest	8306	yes	0.03	0.025	
Spotted Flycatcher	1322	yes	0.05	0.14	
European Pied Flycatcher	666	yes	0.05	0.16	
Long-tailed Tit	623	yes	0.07	0.15	
Eurasian Blue Tit	7105	yes	0.03	0.015	
Great Tit	13596	yes	0.03	0.035	
Coal Tit	15799	yes	0.02	0.035	
European Crested Tit	4042	yes	0.04	0.025	
Marsh Tit	2627	yes	0.04	0.01	
Alpine or Willow Tit	197	yes	0.05	0.05	like Alpine Tit
Alpine Tit	2917	yes	0.05	0.05	
Willow Tit	132	yes	0.05	0.05	like Alpine Tit
Eurasian Treecreeper	3362	yes	0.05	0.055	
Short-toed Treecreeper	1503	yes	0.05	0.055	
Eurasian Golden Oriole	153	yes	0.05	0.2	manually to 0.2
Red-backed Shrike	402	yes	0.05	0.34	
Eurasian Jay	2430	yes	0.05	0.49	
Eurasian Magpie	1184	yes	0.05	0.48	
Spotted Nutcracker	860	yes	0.2	0.49	
Alpine Chough	578	yes	0.05	0.47	
Carrion Crow	3693	yes	0.1	0.09	
Hooded Crow	126	yes	0.05	0.09	like Carrion Crow
Carrion x Hooded Crow	1	yes	0.05	0.09	like Carrion Crow
Common Starling	3897	yes	0.05	0.04	
House Sparrow	15292	yes	0.02	0.025	
Italian Sparrow	1239	yes	0.03	0.03	
House x Italian Sparrow	52	yes	0.1	0.16	
Eurasian Tree Sparrow	2388	yes	0.05	0.18	
White-winged Snowfinch	325	yes	0.1	0.5	
Common Chaffinch	34349	yes	0.02	0.035	
European Serin	1299	yes	0.1	0.11	
Citril Finch	423	yes	0.3	0.17	
European Greenfinch	2910	yes	0.1	0.085	
European Goldfinch	1820	yes	0.2	0.49	
Common Linnet	1048	yes	0.25	0.46	
Lesser Redpoll	628	yes	0.1	0.46	
Red Crossbill	787	yes	0.3	0.46	
Eurasian Bullfinch	1598	yes	0.1	0.26	
Hawfinch	396	yes	0.1	0.38	
Yellowhammer	2212	yes	0.1	0.16	
Cirl Bunting	50	yes	0.05	0.2	manually to 0.2
Rock Bunting	658	yes	0.05	0.15	
Greylag Goose	4	no	0.05	0.05	
Tufted Duck	1	no	0.05	0.05	
Common Merganser	10	no	0.05	0.05	
Common Quail	23	no	0.05	0.05	
Rock Partridge	72	no	0.05	0.05	
Western Capercaillie	8	no	0.05	0.05	
Little Grebe	35	no	0.05	0.05	
Great Cormorant	0	no	0.05	0.05	
Grey Heron	37	no	0.05	0.05	
White Stork	9	no	0.05	0.05	
European Honey Buzzard	32	no	0.25	0.25	
Black Kite	441	no	0.8	0.8	

Northern Goshawk	55	no	0.25	0.25
Eurasian Sparrowhawk	121	no	0.25	0.25
Golden Eagle	5	no	0.25	0.25
Eurasian Hobby	33	no	0.25	0.25
Peregrine Falcon	7	no	0.25	0.25
Water Rail	3	no	0.05	0.05
Corn Crake	1	no	0.05	0.05
Northern Lapwing	9	no	0.05	0.05
Little Ringed Plover	3	no	0.05	0.05
Common Sandpiper	2	no	0.05	0.05
Eurasian Woodcock	18	no	0.05	0.05
Yellow-legged Gull	13	no	0.05	0.05
Stock Dove	132	no	0.05	0.05
European Turtle Dove	29	no	0.05	0.05
Common Cuckoo	924	no	0.5	0.5
Eurasian Scops Owl	1	no	0.05	0.05
Eurasian Pygmy Owl	12	no	0.05	0.05
Tawny Owl	135	no	0.05	0.05
Long-eared Owl	10	no	0.05	0.05
Boreal Owl	7	no	0.05	0.05
European Nightjar	0	no	0.05	0.05
Alpine Swift	53	no	0.05	0.05
Common Kingfisher	26	no	0.05	0.05
Eurasian Hoopoe	4	no	0.05	0.05
Grey-headed Woodpecker	24	no	0.05	0.05
Black Woodpecker	627	no	0.7	0.7
Middle Spotted Woodpecker	75	no	0.05	0.05
White-backed Woodpecker	1	no	0.05	0.05
Eurasian Three-toed Woodpecker	60	no	0.05	0.05
Tawny Pipit	2	no	0.05	0.05
Meadow Pipit	27	no	0.05	0.05
Western Yellow Wagtail	6	no	0.05	0.05
White-throated Dipper	190	no	0.05	0.05
European Stonechat	43	no	0.05	0.05
Cetti's Warbler	1	no	0.05	0.05
Common Grasshopper Warbler	5	no	0.05	0.05
Savi's Warbler	5	no	0.05	0.05
Icterine Warbler	11	no	0.05	0.05
Melodious Warbler	12	no	0.05	0.05
Marsh Warbler	169	no	0.05	0.05
Great Reed Warbler	4	no	0.05	0.05
Common Whitethroat	46	no	0.05	0.05
Collared Flycatcher	10	no	0.05	0.05
Bearded Reedling	1	no	0.05	0.05
Eurasian Nuthatch	3987	no	0.03	0.03
Wallcreeper	38	no	0.05	0.05
Red-billed Chough	30	no	0.05	0.05
Western Jackdaw	32	no	0.05	0.05
Rook	71	no	0.05	0.05
Northern Raven	429	no	0.25	0.25
Eurasian Siskin	304	no	0.2	0.2
Common Rosefinch	5	no	0.05	0.05
Ortolan Bunting	0	no	0.05	0.05
Common Reed Bunting	33	no	0.05	0.05
Corn Bunting	8	no	0.05	0.05

7.2 Normalisation distance: data-defined vs. expert-defined distances

Table 22: Share of sample sites and territories where territories are delimited by data derived normalisation distances. Reading example: Both the Mean Nearest Neighbour and Quantile Value approach used data derived (instead of predefined normalisation distances) for 99% of the territories. For 7% of the sample sites the Quantile Value approach used however predefined normalisation distances.

species	sample sites	manual territories	Mean Nearest Neighbour		Quantile Value	
			sample sites	territories	sample sites	territories
Common Chaffinch	1137	34349	94%	99%	93%	99%
Eurasian Blackcap	1010	20734	90%	99%	88%	98%
Common Blackbird	1034	19205	80%	96%	73%	93%
European Robin	1043	16134	87%	97%	82%	96%
Coal Tit	951	15799	82%	97%	76%	95%
House Sparrow	502	15292	45%	74%	5%	9%
Eurasian Wren	1085	14378	86%	97%	81%	95%
Great Tit	904	13596	81%	96%	78%	95%
Song Thrush	1001	9766	69%	90%	60%	84%
Black Redstart	1157	9209	66%	86%	55%	79%
Common Firecrest	806	8306	74%	94%	63%	90%
Common Chiffchaff	995	8274	72%	91%	61%	86%
Eurasian Blue Tit	767	7105	55%	81%	49%	78%
Dunnock	809	6926	68%	91%	57%	86%
Goldcrest	816	6647	60%	86%	48%	77%
Water Pipit	475	5983	61%	94%	53%	89%
Mistle Thrush	894	5120	36%	58%	28%	51%
Common Wood Pigeon	773	4781	53%	82%	38%	66%
European Crested Tit	698	4042	40%	72%	32%	67%
Eurasian Nuthatch	785	3987	31%	52%	24%	46%
Common Starling	515	3897	35%	58%	23%	44%
Carrion Crow	841	3693	17%	31%	10%	21%
Eurasian Treecreeper	713	3362	42%	69%	27%	56%
White Wagtail	834	3057	15%	30%	7%	16%
Great Spotted Woodpecker	882	3011	23%	45%	7%	15%
Alpine Tit	395	2913	58%	84%	46%	75%
European Greenfinch	555	2910	45%	75%	31%	60%
Tree Pipit	489	2881	67%	92%	54%	85%
Marsh Tit	621	2627	22%	45%	15%	34%
Barn Swallow	528	2472	13%	26%	2%	4%
Eurasian Jay	854	2430	9%	19%	0%	0%
Eurasian Tree Sparrow	370	2388	28%	49%	12%	23%
Yellowhammer	435	2212	55%	80%	24%	37%
Common House Martin	280	2157	25%	52%	10%	30%
Ring Ouzel	399	1939	36%	69%	17%	34%
Western Bonelli's Warbler	305	1891	53%	88%	43%	83%
European Goldfinch	567	1820	27%	52%	5%	10%
Eurasian Bullfinch	567	1598	23%	45%	6%	14%
Northern Wheatear	312	1569	53%	86%	38%	71%
Short-toed Treecreeper	402	1503	35%	63%	22%	48%
Fieldfare	398	1355	12%	25%	2%	5%
Spotted Flycatcher	477	1322	22%	47%	8%	21%
Garden Warbler	435	1305	34%	65%	21%	52%
European Serin	367	1299	37%	70%	22%	51%
Italian Sparrow	38	1239	61%	91%	32%	44%
Eurasian Magpie	430	1184	12%	27%	1%	2%
Common Swift	330	1138	3%	8%	0%	0%
Common Linnet	396	1048	29%	55%	7%	13%

Common Buzzard	731	987	4%	9%	1%	2%
Eurasian Skylark	218	955	55%	86%	37%	68%
Common Cuckoo	598	924	24%	45%	7%	17%
European Green Woodpecker	574	918	16%	31%	4%	10%
Spotted Nutcracker	328	860	13%	39%	1%	3%
Alpine Accentor	217	848	38%	64%	4%	6%
Lesser Whitethroat	293	825	32%	55%	10%	17%
Red Crossbill	354	787	8%	21%	1%	3%
Feral Pigeon	78	773	9%	44%	3%	28%
Common Redstart	265	762	35%	66%	10%	20%
Mallard	289	714	27%	59%	10%	33%
European Pied Flycatcher	277	666	25%	53%	7%	17%
Rock Bunting	149	658	39%	66%	20%	39%
Common Kestrel	606	656	7%	16%	0%	1%
Black Grouse	235	651	28%	52%	5%	15%
Lesser Redpoll	249	628	22%	45%	3%	6%
Black Woodpecker	519	627	7%	14%	1%	4%
Long-tailed Tit	328	622	5%	12%	1%	4%
Alpine Chough	257	578	5%	11%	1%	2%
Grey Wagtail	358	531	11%	24%	3%	6%
Red Kite	455	510	4%	12%	1%	4%
Eurasian Collared Dove	137	450	37%	72%	15%	30%
Whinchat	108	443	52%	82%	33%	56%
Black Kite	383	441	4%	13%	0%	0%
Northern Raven	443	429	5%	11%	0%	1%
Citril Finch	178	423	22%	44%	7%	19%
Red-backed Shrike	203	402	23%	44%	4%	9%
Hawfinch	208	396	4%	15%	0%	6%
White-winged Snowfinch	118	325	14%	28%	3%	5%
Eurasian Siskin	149	304	9%	22%	1%	7%
Eurasian Reed Warbler	50	290	50%	90%	34%	74%
Eurasian Crag Martin	160	287	5%	12%	1%	1%
Eurasian Coot	70	266	29%	52%	20%	39%
Wood Warbler	145	243	19%	45%	3%	7%
Rock Ptarmigan	142	234	15%	26%	1%	2%
Alpine or Willow Tit	48	197	27%	73%	19%	63%
White-throated Dipper	166	190	2%	5%	0%	0%
Marsh Warbler	58	169	40%	76%	28%	62%
Eurasian Golden Oriole	80	153	19%	45%	9%	28%
Tawny Owl	117	135	4%	8%	0%	0%
Stock Dove	100	132	7%	11%	2%	5%
Willow Tit	63	131	19%	39%	2%	6%
Hooded Crow	50	126	28%	53%	6%	10%
Eurasian Sparrowhawk	122	121	1%	2%	0%	0%
Willow Warbler	86	113	28%	52%	5%	10%
Eurasian Wryneck	68	96	31%	52%	6%	13%
Common Nightingale	37	90	32%	58%	19%	42%
Hazel Grouse	68	88	1%	8%	1%	7%
Common Rock Thrush	65	80	12%	22%	0%	0%
Middle Spotted Woodpecker	43	75	0%	0%	0%	0%
Rock Partridge	57	72	9%	19%	2%	8%
Rook	6	71	17%	52%	17%	52%
Eurasian Three-toed Woodpecker	49	60	6%	17%	2%	14%
Great Crested Grebe	22	59	45%	80%	14%	18%
Northern Goshawk	55	55	2%	4%	0%	0%
Alpine Swift	6	53	0%	0%	0%	0%
Lesser Spotted Woodpecker	47	53	2%	4%	2%	4%
House x Italian Sparrow	8	52	25%	63%	0%	0%
Cirl Bunting	24	50	25%	41%	8%	20%
Common Whitethroat	36	46	11%	27%	6%	16%

|						
European Stonechat	25	43	20%	24%	4%	6%
Wallcreeper	37	38	0%	0%	0%	0%
Grey Heron	14	37	7%	6%	0%	0%
Mute Swan	29	35	7%	6%	3%	2%
Little Grebe	19	35	11%	15%	0%	0%
Eurasian Hobby	34	33	6%	9%	0%	0%
Common Reed Bunting	21	33	29%	50%	5%	13%
European Honey Buzzard	33	32	3%	6%	0%	0%
Western Jackdaw	10	32	0%	0%	0%	0%
Red-billed Chough	21	30	0%	0%	0%	0%
European Turtle Dove	16	29	38%	63%	25%	53%
Meadow Pipit	29	27	17%	40%	3%	14%
Common Kingfisher	23	26	0%	0%	0%	0%
Grey-headed Woodpecker	19	24	11%	28%	0%	0%
Common Moorhen	16	24	6%	18%	6%	18%
Common Quail	20	23	5%	9%	0%	0%
Woodlark	9	21	22%	52%	11%	30%
Eurasian Woodcock	14	18	21%	39%	0%	0%
Common Pheasant	4	15	50%	72%	0%	0%
Yellow-legged Gull	13	13	0%	0%	0%	0%
Eurasian Pygmy Owl	12	12	0%	0%	0%	0%
Melodious Warbler	10	12	10%	23%	0%	0%
Icterine Warbler	11	11	0%	0%	0%	0%
Long-eared Owl	10	10	0%	0%	0%	0%
Collared Flycatcher	4	10	75%	80%	25%	40%
Common Merganser	8	10	63%	67%	13%	11%
White Stork	5	9	40%	67%	40%	67%
Northern Lapwing	1	9	0%	0%	0%	0%
Corn Bunting	4	8	75%	92%	50%	75%
Western Capercaillie	8	8	0%	0%	0%	0%
Peregrine Falcon	22	7	0%	0%	0%	0%
Boreal Owl	7	7	0%	0%	0%	0%
Western Yellow Wagtail	4	6	25%	25%	25%	25%
Common Grasshopper Warbler	9	5	11%	27%	11%	27%
Golden Eagle	6	5	0%	0%	0%	0%
Savi's Warbler	3	5	33%	40%	0%	0%
Common Rosefinch	4	5	0%	0%	0%	0%
Greylag Goose	4	4	0%	0%	0%	0%
Great Reed Warbler	4	4	0%	0%	0%	0%
Eurasian Hoopoe	3	4	0%	0%	0%	0%
Water Rail	3	3	0%	0%	0%	0%
Little Ringed Plover	1	3	0%	0%	0%	0%
Tawny Pipit	2	2	0%	0%	0%	0%
Tufted Duck	1	1	100%	100%	0%	0%
Carrion x Hooded Crow	2	1	0%	0%	0%	0%
Corn Crake	1	1	0%	0%	0%	0%
Eurasian Scops Owl	1	1	0%	0%	0%	0%
White-backed Woodpecker	1	1	0%	0%	0%	0%
Bearded Reedling	1	1	0%	0%	0%	0%
Cetti's Warbler	1	1	0%	0%	0%	0%
European Nightjar	1	0	0%	0%	0%	0%
Ortolan Bunting	1	0	0%	0%	0%	0%

7.3 Accuracy values for all species

Table 23: Accuracy values per species for the three different approaches for identifying the normalisation distance. The list is sorted according to the manually delimited number of territories in MHB and BDM sample sites analysed from 2013–2016.

Species	Count manual	Linear Regression MAE_{sp}	Quantile Value MAE_{sp}	Mean Nearest Neighbour MAE_{sp}	Linear Regression ME_{sp}	Quantile Value ME_{sp}	Mean Nearest Neighbour ME_{sp}
Common Chaffinch	34349	0.164	0.083	0.076	-0.075	0.004	0.008
Eurasian Blackcap	20734	0.153	0.078	0.071	-0.054	0.019	-0.017
Common Blackbird	19205	0.163	0.113	0.097	-0.007	0.018	0.007
European Robin	16134	0.145	0.093	0.085	0.008	0.018	-0.004
Coal Tit	15799	0.155	0.095	0.085	-0.028	0.015	-0.013
House Sparrow	15292	0.192	0.175	0.203	-0.105	-0.082	-0.131
Eurasian Wren	14378	0.155	0.084	0.076	0.035	0.019	-0.003
Great Tit	13596	0.149	0.112	0.101	-0.013	0.042	0.001
Song Thrush	9766	0.132	0.096	0.096	0.057	0.017	0.052
Black Redstart	9209	0.151	0.103	0.092	0.093	0.034	0.038
Common Firecrest	8306	0.153	0.088	0.091	-0.052	0.016	-0.067
Common Chiffchaff	8274	0.149	0.092	0.084	0.063	0.019	0.010
Eurasian Blue Tit	7105	0.144	0.163	0.112	-0.003	0.061	-0.019
Dunnock	6926	0.130	0.084	0.074	0.048	0.001	0.008
Goldcrest	6647	0.118	0.075	0.070	-0.023	0.008	-0.016
Water Pipit	5983	0.175	0.140	0.131	0.085	0.008	0.098
Mistle Thrush	5120	0.207	0.179	0.177	0.175	0.122	0.132
Common Wood Pigeon	4781	0.248	0.172	0.172	0.207	-0.027	0.107
European Crested Tit	4042	0.132	0.129	0.100	0.011	0.007	-0.017
Eurasian Nuthatch	3987	0.132	0.126	0.114	0.031	-0.023	-0.014
Common Starling	3897	0.238	0.220	0.201	0.150	0.050	0.030
Carrion Crow	3693	0.367	0.264	0.278	0.353	0.226	0.240
Eurasian Treecreeper	3362	0.139	0.125	0.103	0.060	0.006	0.024
White Wagtail	3057	0.177	0.164	0.149	0.114	0.007	0.000
Great Spotted Woodpecker	3011	0.274	0.147	0.207	0.234	0.000	0.103
Alpine Tit	2913	0.137	0.117	0.101	0.041	0.041	0.039
European Greenfinch	2910	0.155	0.114	0.109	0.105	0.011	0.021
Tree Pipit	2881	0.181	0.115	0.105	0.134	0.002	0.042
Marsh Tit	2627	0.104	0.104	0.098	0.057	0.063	0.063
Barn Swallow	2472	0.251	0.208	0.212	0.172	0.063	0.079
Eurasian Jay	2430	0.253	0.150	0.186	0.203	0.040	0.094
Eurasian Tree Sparrow	2388	0.153	0.182	0.166	0.062	-0.097	-0.058
Yellowhammer	2212	0.177	0.104	0.136	0.153	0.000	0.102
Common House Martin	2157	0.245	0.203	0.217	0.132	0.036	0.057
Ring Ouzel	1939	0.237	0.159	0.171	0.219	0.041	0.137
Western Bonelli's Warbler	1891	0.120	0.072	0.069	0.085	0.009	0.033
European Goldfinch	1820	0.232	0.161	0.172	0.169	-0.006	0.086
Eurasian Bullfinch	1598	0.162	0.128	0.129	0.100	-0.049	0.036
Northern Wheatear	1569	0.226	0.202	0.157	0.196	-0.047	0.088
Short-toed Treecreeper	1503	0.140	0.117	0.107	0.078	0.022	0.043
Fieldfare	1355	0.295	0.185	0.198	0.270	0.063	0.117
Spotted Flycatcher	1322	0.077	0.082	0.061	0.014	-0.048	0.006
Garden Warbler	1305	0.089	0.068	0.042	0.027	-0.038	-0.002
European Serin	1299	0.158	0.106	0.108	0.124	0.017	0.076
Italian Sparrow	1239	0.160	0.143	0.140	0.088	0.055	-0.002
Eurasian Magpie	1184	0.335	0.163	0.211	0.308	0.067	0.142
Common Swift	1138	0.322	0.295	0.304	0.153	0.093	0.108
Common Linnet	1048	0.232	0.203	0.185	0.167	-0.055	0.058
Common Buzzard	987	0.293	0.217	0.229	0.243	0.157	0.170

Eurasian Skylark	955	0.247	0.124	0.128	0.231	-0.032	0.095
Common Cuckoo	924	0.295	0.143	0.225	0.260	0.078	0.190
European Green Woodpecker	918	0.223	0.129	0.170	0.186	0.072	0.129
Spotted Nutcracker	860	0.359	0.166	0.275	0.343	0.091	0.230
Alpine Accentor	848	0.250	0.133	0.183	0.233	0.059	0.154
Lesser Whitethroat	825	0.112	0.076	0.083	0.073	-0.011	0.049
Red Crossbill	787	0.331	0.252	0.283	0.260	0.117	0.165
Feral Pigeon	773	0.391	0.414	0.351	0.354	0.273	0.117
Common Redstart	762	0.113	0.077	0.072	0.092	-0.028	0.055
Mallard	714	0.287	0.254	0.272	0.177	0.092	0.136
European Pied Flycatcher	666	0.173	0.176	0.152	0.130	0.057	0.127
Rock Bunting	658	0.177	0.100	0.107	0.138	-0.035	0.039
Common Kestrel	656	0.155	0.104	0.148	0.132	0.076	0.121
Black Grouse	651	0.217	0.136	0.158	0.167	0.007	0.085
Lesser Redpoll	628	0.233	0.142	0.182	0.187	0.050	0.120
Black Woodpecker	627	0.195	0.127	0.140	0.139	0.066	0.084
Long-tailed Tit	622	0.121	0.080	0.091	0.060	0.009	0.024
Alpine Chough	578	0.483	0.436	0.446	0.251	0.149	0.171
Grey Wagtail	531	0.184	0.114	0.138	0.136	0.023	0.081
Red Kite	510	0.293	0.224	0.247	0.264	0.193	0.217
Eurasian Collared Dove	450	0.306	0.217	0.171	0.265	-0.096	0.061
Whinchat	443	0.193	0.105	0.100	0.148	-0.011	0.062
Black Kite	441	0.265	0.194	0.233	0.235	0.148	0.201
Northern Raven	429	0.224	0.158	0.182	0.213	0.146	0.171
Citril Finch	423	0.212	0.170	0.183	0.126	0.011	0.076
Red-backed Shrike	402	0.149	0.082	0.073	0.119	-0.014	0.044
Hawfinch	396	0.231	0.180	0.188	0.204	0.146	0.155
White-winged Snowfinch	325	0.298	0.216	0.245	0.267	0.155	0.197
Eurasian Siskin	304	0.259	0.196	0.198	0.137	0.052	0.055
Eurasian Reed Warbler	290	0.153	0.123	0.151	-0.006	-0.043	-0.090
Eurasian Crag Martin	287	0.252	0.178	0.201	0.225	0.124	0.149
Eurasian Coot	266	0.309	0.216	0.188	0.277	-0.045	0.027
Wood Warbler	243	0.176	0.154	0.173	0.142	0.066	0.145
Rock Ptarmigan	234	0.204	0.136	0.170	0.176	0.098	0.141
Alpine or Willow Tit	197	0.126	0.097	0.082	0.042	0.000	0.024
White-throated Dipper	190	0.084	0.071	0.080	0.044	0.010	0.020
Marsh Warbler	169	0.035	0.023	0.041	0.012	0.000	-0.029
Eurasian Golden Oriole	153	0.193	0.146	0.152	0.171	0.064	0.129
Tawny Owl	135	0.140	0.069	0.088	0.140	0.069	0.088
Stock Dove	132	0.280	0.268	0.272	0.269	0.257	0.261
Willow Tit	131	0.272	0.311	0.333	-0.213	-0.252	-0.274
Hooded Crow	126	0.316	0.317	0.331	0.236	0.118	0.000
Eurasian Sparrowhawk	121	0.008	0.008	0.016	0.008	0.008	0.016
Willow Warbler	113	0.368	0.346	0.339	0.310	0.259	0.291
Eurasian Wryneck	96	0.234	0.159	0.228	0.218	0.142	0.211
Common Nightingale	90	0.074	0.064	0.043	0.032	0.021	0.022
Hazel Grouse	88	0.074	0.054	0.065	0.053	0.033	0.043
Common Rock Thrush	80	0.223	0.184	0.200	0.223	0.184	0.200
Middle Spotted Woodpecker	75	0.194	0.039	0.051	0.194	0.013	0.051
Rock Partridge	72	0.148	0.150	0.138	0.074	0.050	0.063
Rook	71	0.289	0.111	0.111	0.133	-0.083	-0.083
Eurasian Three-toed Woodpecker	60	0.200	0.178	0.155	0.200	0.178	0.155
Great Crested Grebe	59	0.514	0.380	0.449	0.423	-0.042	0.225
Northern Goshawk	55	0.000	0.000	0.000	0.000	0.000	0.000
Alpine Swift	53	0.190	0.190	0.190	0.127	0.127	0.127
Lesser Spotted Woodpecker	53	0.086	0.070	0.070	0.086	0.070	0.070
House x Italian Sparrow	52	0.365	0.365	0.423	-0.365	-0.365	-0.423
Cirl Bunting	50	0.169	0.143	0.140	0.136	0.071	0.105
Common Whitethroat	46	0.098	0.061	0.061	0.098	0.061	0.061
European Stonechat	43	0.233	0.256	0.233	-0.233	-0.256	-0.233

Wallcreeper	38	0.050	0.050	0.050	0.050	0.050	0.050
Grey Heron	37	0.154	0.154	0.154	-0.051	-0.051	-0.051
Little Grebe	35	0.158	0.167	0.111	0.000	-0.111	-0.056
Mute Swan	35	0.849	0.849	0.851	0.849	0.831	0.833
Eurasian Hobby	33	0.029	0.029	0.029	0.029	0.029	0.029
Common Reed Bunting	33	0.114	0.088	0.029	0.000	-0.029	0.029
European Honey Buzzard	32	0.059	0.059	0.059	0.059	0.059	0.059
Western Jackdaw	32	0.610	0.538	0.538	-0.171	-0.179	-0.179
Red-billed Chough	30	0.370	0.356	0.356	0.326	0.311	0.311
European Turtle Dove	29	0.212	0.212	0.156	0.030	0.030	0.031
Meadow Pipit	27	0.612	0.510	0.509	0.582	0.431	0.472
Common Kingfisher	26	0.161	0.161	0.161	0.161	0.161	0.161
Common Moorhen	24	0.148	0.083	0.083	0.074	-0.083	-0.083
Grey-headed Woodpecker	24	0.333	0.294	0.250	0.333	0.294	0.250
Common Quail	23	0.083	0.083	0.083	0.000	0.000	0.000
Woodlark	21	0.091	0.048	0.087	0.000	-0.048	0.087
Eurasian Woodcock	18	0.286	0.250	0.200	0.000	-0.050	0.000
Common Pheasant	15	0.167	0.063	0.167	0.167	0.063	0.167
Yellow-legged Gull	13	0.519	0.519	0.519	0.519	0.519	0.519
Eurasian Pygmy Owl	12	0.077	0.077	0.077	0.077	0.077	0.077
Melodious Warbler	12	0.077	0.077	0.077	0.077	0.077	0.077
Icterine Warbler	11	0.214	0.214	0.214	0.214	0.214	0.214
Common Merganser	10	0.100	0.100	0.100	-0.100	-0.100	-0.100
Long-eared Owl	10	0.000	0.000	0.000	0.000	0.000	0.000
Collared Flycatcher	10	0.182	0.182	0.182	0.000	0.000	0.000
White Stork	9	0.000	0.000	0.000	0.000	0.000	0.000
Northern Lapwing	9	0.889	0.889	0.889	-0.889	-0.889	-0.889
Western Capercaillie	8	0.000	0.000	0.000	0.000	0.000	0.000
Corn Bunting	8	0.467	0.333	0.385	0.467	0.333	0.385
Peregrine Falcon	7	0.682	0.682	0.682	0.682	0.682	0.682
Boreal Owl	7	0.000	0.000	0.000	0.000	0.000	0.000
Western Yellow Wagtail	6	0.333	0.333	0.333	-0.333	-0.333	-0.333
Golden Eagle	5	0.167	0.167	0.167	0.167	0.167	0.167
Common Grasshopper Warbler	5	0.545	0.545	0.545	0.545	0.545	0.545
Savi's Warbler	5	0.000	0.000	0.000	0.000	0.000	0.000
Common Rosefinch	5	0.000	0.000	0.000	0.000	0.000	0.000
Greylag Goose	4	0.636	0.667	0.667	0.636	0.667	0.667
Eurasian Hoopoe	4	0.250	0.250	0.250	-0.250	-0.250	-0.250
Great Reed Warbler	4	0.000	0.000	0.000	0.000	0.000	0.000
Water Rail	3	0.000	0.000	0.000	0.000	0.000	0.000
Little Ringed Plover	3	0.667	0.667	0.667	-0.667	-0.667	-0.667
Common Sandpiper	2	1.000	1.000	1.000	-1.000	-1.000	-1.000
Tawny Pipit	2	0.333	0.333	0.333	0.333	0.333	0.333
Tufted Duck	1	0.667	0.667	0.667	0.667	0.667	0.667
Corn Crake	1	0.000	0.000	0.000	0.000	0.000	0.000
Eurasian Scops Owl	1	0.000	0.000	0.000	0.000	0.000	0.000
White-backed Woodpecker	1	0.000	0.000	0.000	0.000	0.000	0.000
Carrion x Hooded Crow	1	0.500	0.500	0.500	0.500	0.500	0.500
Bearded Reedling	1	0.000	0.000	0.000	0.000	0.000	0.000
Cetti's Warbler	1	0.000	0.000	0.000	0.000	0.000	0.000
Great Cormorant	0	NA	NA	NA	NA	NA	NA
European Nightjar	0	1.000	1.000	1.000	1.000	1.000	1.000
Ortolan Bunting	0	1.000	1.000	1.000	1.000	1.000	1.000

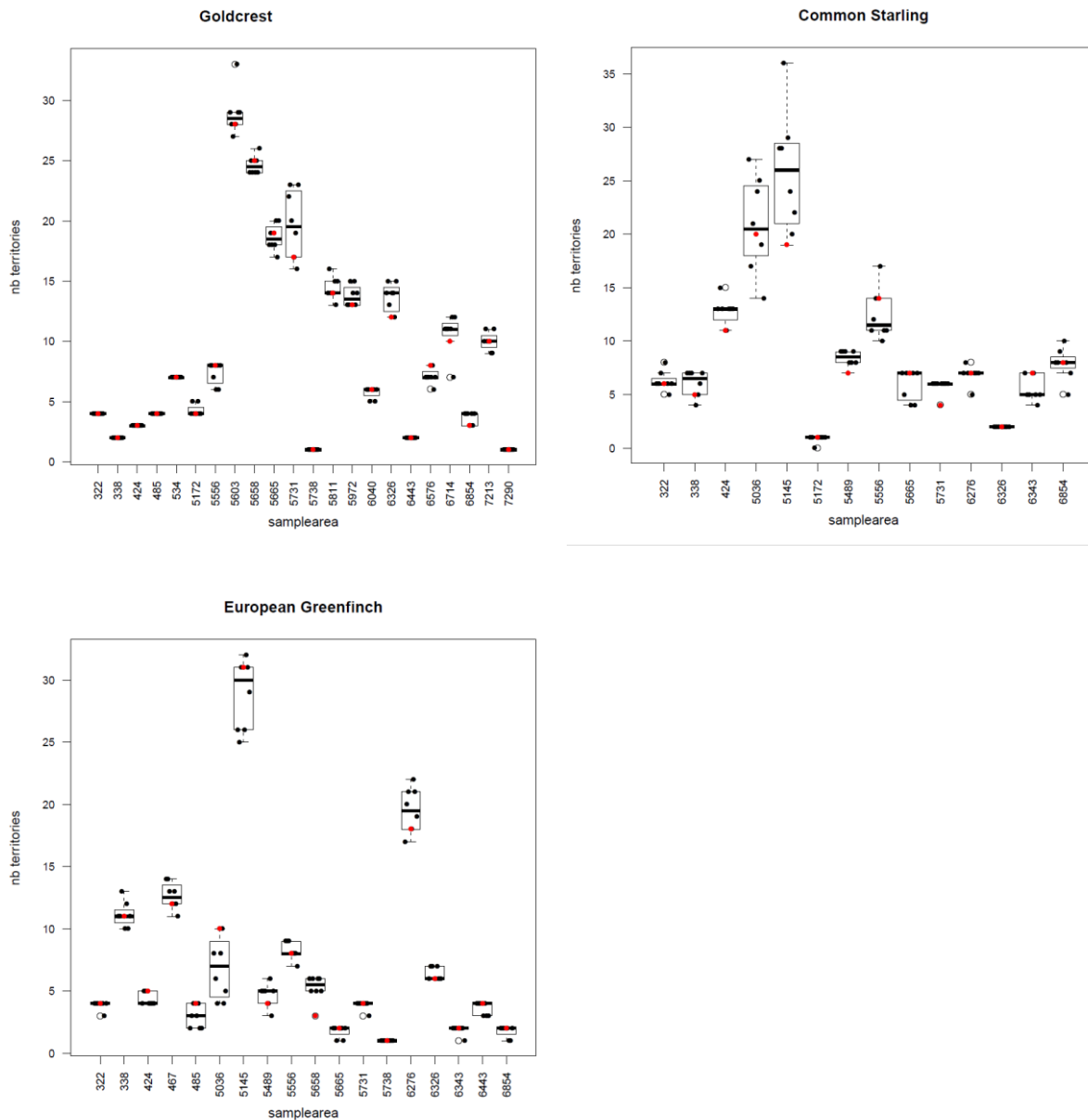## 7.4 Variation in manual territory delimitation by experts
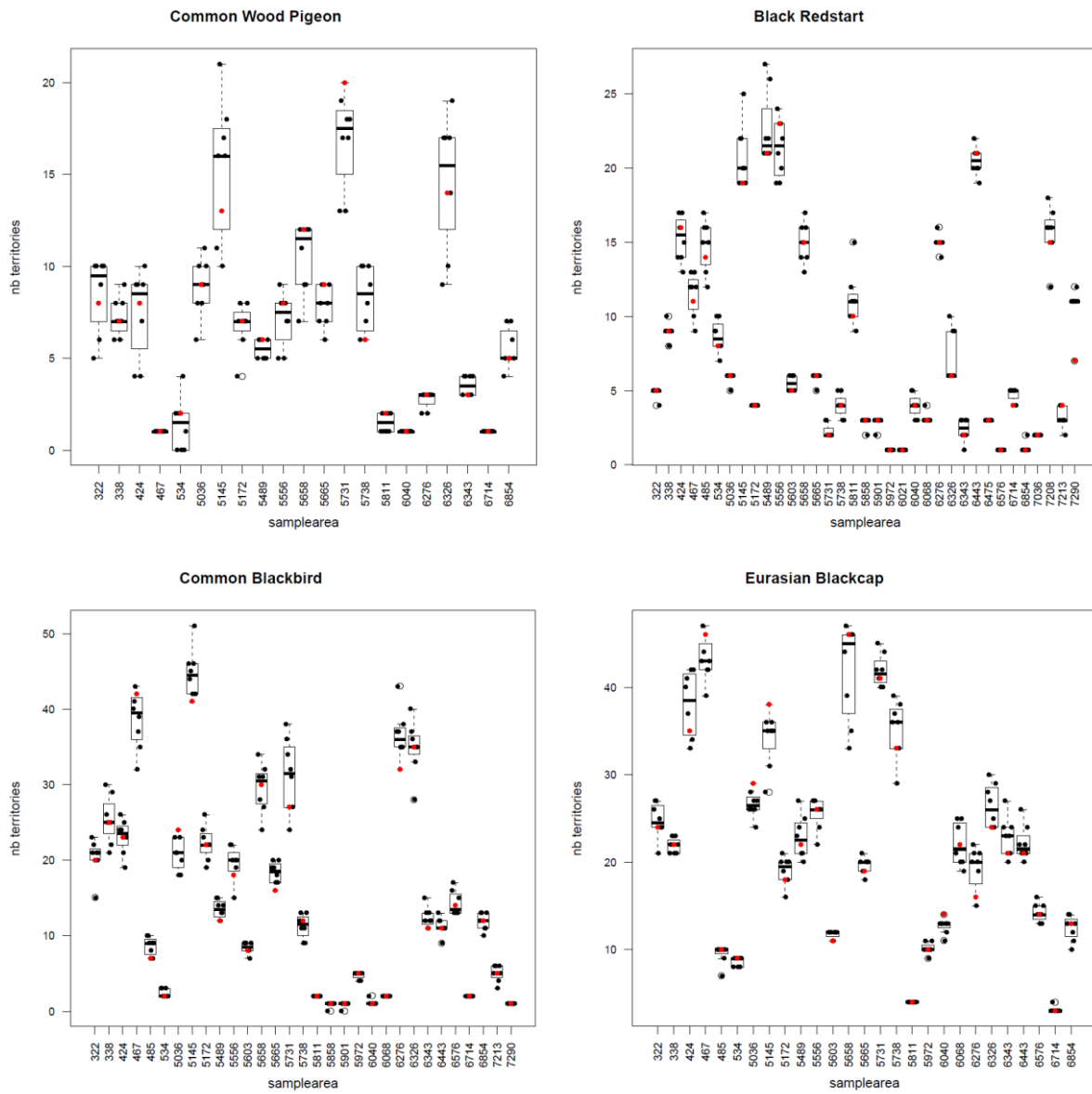
Figure 35: Territory delimitation for seven species by eight different experts. Boxplots show the range of eight different manual territory delimitations for different sample sites. Dots show the count values that are behind the boxplots. Red dots show values of the "original user" (i.e. the person also conducting the field work) after validation by the Swiss Ornithological Institute.

7.5 Species-specific coefficient of variation (CV)

Table 24: Species-specific coefficient of variations (CV) for the analyst effect, i.e. the differences that appear when different people delimit territories with the same data.

species	this study	O'Conner & Marchant (1981)	Svensson (1974)	Bötsch (unpublished data)
Common Wood Pigeon	20.3%	14.7%		
Eurasian Collared Dove		19.9%		
Tree Pipit		12.0%		
Eurasian Wren		6.3%		
Dunnock		3.4%		
European Robin		3.2%		
Black Redstart	10.3%			
Common Blackbird	10.5%	3.0%	36.0%	
Song Thrush		3.9%		
Mistle Thrush		23.3%		
Common Whitethroat		8.7%		
Eurasian Blackcap	8.2%	7.5%		
Willow Warbler		4.5%	16.0%	
Goldcrest	7.2%	9.3%		
Long-tailed Tit		9.5%		
Eurasian Blue Tit		5.1%		13.3%
Great Tit		10.2%		20.3%
Coal Tit		3.1%		
Marsh Tit		12.6%		
Eurasian Magpie		18.8%		
Common Starling	17.5%			
Common Chaffinch		3.3%		
European Greenfinch	12.6%	18.0%		
European Goldfinch		12.0%		
Common Linnet		22.1%		
Eurasian Bullfinch		11.5%		
Yellowhammer		11.1%		