IV.B.5 Discussion

Most of the above findings coincide with the results of testing the performance and accuracy of GPS collars in boreal forests (Rempel et al. 1995, Moen et al. 1996, Edenius 1997, Merrill et al. 1998, Dussault et al. 1999, Bowman et al. 2000).

According to our model the absolute increase in observations when including the 2D observations, was independent of the habitat type. This signifies that the absolute probability of obtaining a 2D observation was independent of the habitat type. However, since the number of 3D observations is higher on the open test-site compared to under canopy, the ratio 2D/3D is higher under canopy than on open sites. These results are similar to the results of previous tests (Rempel et al. 1995, Moen et al. 1997, Rempel and Rodgers 1997).
The opinions on the exact cause (tree height, basal area or canopy cover) of the reduction in observation probability are contradicting. The narrow range of tested vegetation characteristics in each experiment (Rempel et al. 1995, Dussault et al. 1999, Bowman et al. 2000), the differences between the study areas and the lack of repeated tests for each test-site, probably explain the contradicting results. The observed higher success-rate for the coniferous test-sites, in our study, could be related to the fact that the average tree height was higher for the deciduous test-sites, thereby confirming the hypothesis that GPS performance decreases as trees get taller (Rempel et al. 1995, Rodgers et al. 1996, Dussault et al. 1999). In practice factors such as stem density, tree height and canopy cover can be expected to be highly interrelated (Edenius 1997), making the discussion above rather theoretical. Summarising we can conclude that our results of testing the performance of a GPS

collar in temperate forests are similar to previous tests in boreal forests. The low (compared to previous published results) overall success-rate in the deciduous stands illustrates the negative influence of the height of the trees, typically for old, deciduous temperate forest stands.

Some authors found a negative influence of animal movement on the probability of obtaining a location (Edenius 1997, Bowman et al. 1999) but not of the position of the head; others discovered the opposite (Moen et al. 1996). Our results clearly indicate a greater observation probability when the collar is in a vertical position (antenna facing the sky) compared to an inclined position. This coincides with the results of the experimental set-up of Moen et al. in 1996. The translation of these results into the influence of animal behaviour and/or head position on the observation rate depends on the method of collar fixation and the antenna position in the collar.

For the open test-site and the deciduous forest stands, the achieved accuracy and the relationship between DOP and location accuracy agree with the results found by most authors testing collars returning uncorrected GPS locations (Rempel et al. 1995, Moen et al. 1996, Edenius 1997, Bowman et al. 2000). The influence of the DOP on the accuracy of the locations is greater for the 2D observations than for the 3D.

The systematic greater location error and even systematic bias found for the coniferous test-sites is, however, exceptional and cannot be explained. Only Rempel et al. found significant higher location errors in a dense, Red pine (*Pinus resinosa),* forest stand, compared to other forest stands in the study area. However this coincided with a significant lower observation rate.

Due to the design of the experiment we can not differ a possible seasonal effect from the observed habitat-related differences in observation rate.

80 % of the observations happened in the first two minutes and only 25 % of the observations obtained in the two last minutes were 3D observations. Consequently, a shorter total search period could be programmed, lowering the total battery drain. However, programming a longer search period after a 2D location is obtained, would probably increase the percentage of 3D observations.

IV.B.6 Conclusions and implications

Because of the influence of the habitat on the probability of obtaining a location, the analysis of habitat preference based on animal fixes gathered by GPS collars risks producing false conclusions. Habitat preference is analysed comparing the relative use of a habitat type with its availability. Given the fact that during the experiment the collar was standing on each of the seven sites for six days, the true use of the seven sites was the same as the availability of each site, being 1/7 of the total time.

Using a Chi-square statistic to test the H_0 that animals use the different sites randomly or proportionally to their occurrence (Manly et al. 1993), the researcher would, falsely, reject the H_0 when using the number of GPS fixes for each of the seven test sites ($\chi^2 = 15.6$, df =6, p= 0.016). Although only very roughly, the model-estimated contrasts in success-rate based on the habitat type allows for the correction of the observed number of fixes. Using these corrected numbers of observations for each of the seven test sites, the H_0 is no longer rejected ($\chi^2 = 4.68$, df = 6, p= 0.58). This clearly shows the possible use of models allowing a correction for habitat-related differences in observation rate when using GPS collars.

Since the time of the observations is known, interpolation between two successful localisation attempts could be a method to generate the missing locations. The use of movement vectors, rather than animal locations, could be another way to approach the problem (Rempel et al. 1995)

In western and central Europe, it is doubtful that the high cost of a non-differential GPS collar with the above-mentioned accuracy can be justified in habitat analysis studies of roe deer, as they are a territorial species using small scale, highly fragmented landscapes. The recently published results on the use of differentially corrected GPS data for wildlife studies (Moen et al. 1997, Rempel and Rodgers 1997) demonstrate the gain in accuracy obtained when applying this technology. Depending on the resolution needed for a particular project, the researcher has to decide if the present accuracy will fulfil the needs. Justifying the extra costs of installation and maintenance of a base station and the post-processing of the collected information also fully depends on the objectives of the study (Rempel and Rodgers 1997).

The further development of automatic telemetry systems using more traditional transmitters and fixed stations has temporarily been abandoned by most wildlife researchers and commercial telemetry system manufactures because of the promising results of the first GPS collars. At present, one could question this decision. Further development of both systems should happen in parallel since both have their own applications, depending on the size of the animals to be studied and geographic extent of the studied process.

Finally the high cost of GPS collars, allows most research organisations to observe only few individuals, and therefore poses the question whether one can infer behaviour patterns of animal populations based on such small number of observed animals (Moen et al. 1996).

Nevertheless, the use op GPS collars allows to record locations both day and night, and the overall cost of collecting many locations on regular time intervals can be dramatically reduced (Merrill et al. 1998). The accuracy of non-corrected GPS locations far exceeds the accuracy of Agros systems (see Taillade 1992) and in some areas even the accuracy of conventional VHF telemetry (see also Obbard et al. 1998). Finally, the use of GPS collars offers greater sampling frequency where sampling by conventional collars is limited by terrain characteristics, availability of personnel, budget or weather constraints (Moen et al. 1996, Rodgers et al. 1996, Obbard et al. 1998)

IV.C Appropriateness of the linear correction method for GPS positional fixes in wildlife studies

IV.C.1 Abstract

This part describes the results of tests performed to evaluate linear correction of GPS measurements as an alternative to differentially correcting GPS positional fixes. Differential correction requires information that is not available with the existing animal-borne GPS systems for smaller mammals. Therefore linear correction, by means of a second GPS rover, has been suggested. To test the accuracy of linearly corrected measurements, we compared the position estimates of raw, linearly corrected and differentially corrected GPS positional fixes versus the true (known) geodetic position. The tests clearly indicate that the accuracy of linear correction is highly unstable and is related to differences in satellite constellation used by the GPS receivers. Linear correction is consequently strongly discouraged, when differential correction is not possible, in favour of raw GPS measurements, of which the error is well known and more predictable.

IV.C.2 Introduction

Although still quite innovative, the application of Global Positioning System (GPS) technology in wildlife research has been documented by several authors over the last couple of years. The weight of collar-borne GPS receivers has, however, limited their use to the tracking of larger animals, and most of it concerns non-differentially corrected GPS positioning. The first experimental results of GPS tracking with post-processing differential correction on moose were published only recently (Moen et al.

1997, Rempel and Rodgers 1997). Differential correction involves using a second GPS (base station) on a known position to collect satellite information from all visible satellites. The information is used to remove the errors in the data collected by the GPS operating in the field. The errors are due to Selective Availability, atmospheric interference, satellite ephemeris errors and clock errors. A requirement to enable differential correction is that the satellite information collected by the base station covers all the satellites used to calculate the position of the GPS in the field and that the distance between the two GPS units does not exceed 500 km.

Such differential correction requires more information than is stored by the first-generation animal-borne GPS collars (Rempel et al. 1995, Moen et al. 1996), necessitating more complex collars, higher data storage capacities, higher battery capacity and thus increased collar weight. Miniaturisation is the natural solution, especially where the collaring of smaller mammals is concerned. While receivers for this group are under development, they will not allow, as yet, differential correction. In the meantime a linear correction method has been proposed by some wildlife researchers and collar manufacturers (Moore et al. 1997). The method encompasses the correction of animal borne GPS-derived positions with simultaneously acquired positional errors of a known stationary location. We tested the accuracy of this linear correction method as an alternative to differentially corrected GPS data. This chapter aims to describe the accuracy of linear correction, to find out if and why there is a difference between linear and differential correction and to evaluate the use of linear correction as a GPS correction method.

IV.C.3 Methods

We collected during 12 data sessions at least 30 positional fixes on two geodetic points using a Trimble ProXL GPS receiver. To process the datafiles we used PFINDER V.3.0 software. The operational parameters were a time interval of 15 minutes between sessions, a logging rate of 5 seconds (interval between individual fixes) and a maximum positional dilution of precision (PDOP) of 6. The GPS was in manual 3D position fix mode, meaning that the elevation was computed by the GPS itself and measurements only occurred if at least 4 satellites were available. Based on the spread of the raw data it is evident that Selective Availability was active during the data collection sessions. One of the geodetic points was situated in the open field, the other in a forested area. The experiment was carried out during winter with only woody matter (stems and branches) obstructing the reception of the GPS satellite signals. Simultaneously acquired data from a base station, located less than 20 kilometres away, were used to differentially correct the locational fixes for the two points. The base station's own position, computed every 5 seconds using the combination of the 4 satellites that resulted in the best PDOP, was used to linearly correct the field data. We first computed for each observation of the base station the difference between the measured and the known northing (respectively easting) and subsequently substracted these differences from the raw field GPS data (northing respectively easting). We then compared the locational error (LE) of the raw, the linearly corrected and the differentially corrected measurements (three correction methods) according to the formula of Rempel et al. (1995). They defined the LE as the Euclidean distance between the estimated and the true location of the fixes.

As an estimator of accuracy, the median of the LE's was calculated for each location and correction method, and this for each data session separately and for all sessions combined. We used the median as a descriptive statistic instead of the mean because of the non-normal distribution of the data. Since the Friedman test is known for its robust non-parametric capabilities, allowing for multiple comparisons between groups using average ranks (Siegel & Castellan 1988), we selected it to test the statistical significance of the observed differences in the LE's obtained for each of the locations with the different correction methods.

To evaluate the effectiveness of the linear and differential correction methods we then computed the estimator (EF). EF portrays the estimation of the accuracy (median) for the linearly and differentially corrected measurements as a percentage of the estimation of the accuracy (median) of the raw measurements

$$EF = 100 * \frac{median_{.c}}{median_{nc}}$$

where $median_{.c}$= median of the LE's of the corrected measurements with

'.' = l for linear correction and

'.' = d for differential correction

$median_{nc}$= median of the LE's of the raw measurements.

We used the Wilcoxon Matched-Pairs rank test to compare the EF's of the linear and differential correction method (Siegel & Castellan 1988).

IV.C.4 Results

Table IV.5 summarises the median values of the LE's for each data sample. The Friedman test clearly illustrates the existence of statistically significant differences between the LE values for the three methods (no correction, linear correction and differential correction), as well with respect to the open field ($p < 0.01$) as to forested conditions in winter ($p = 0.01$). In both cases the LE of the differentially corrected measurements is significantly smaller than the LE of the raw measurements whereas there is no statistically significant difference in LE between raw measurements and linearly corrected measurements.

Table IV.5: Values of the median of the positional errors for the different datasessions, in meters from the true position, (sample size equals or exceeds 30).

	Session	Non corrected	Linearly corrected	Differ. Corrected
Open area	1	37.67	43.85	0.56
	2	29.91	11.92	0.87
	3	33.60	15.00	0.76
	4	31.36	29.90	0.08
	5	13.34	6.69	0.55
	6	8.70	4.49	0.60
	All	26.98	12.88	0.61
Forested area	1	59.52	49.91	9.94
	2	27.72	22.87	2.77
	3	13.83	22.5	4.54
	4	28.25	41.96	6.07
	5	34.87	22.99	8.41
	6	43.13	50.24	5.10
	All	38.90	29.87	4.85

The EF, as previously defined, can easily be calculated from table IV.4, being the third (linear correction) respectively the fourth (differential correction) column of table IV.4 divided by the second column (raw measurements). Multiplying this value with 100 expresses the EF as a percentage. The EF or correction effectiveness varied for the linear correction between 0.39 and 1.16 for the open area and between 0.65 and 1.62 for the forested area. This shows that linear correction might both reduce and increase the LE by as much as 60%. The EF for differentially corrected measurements never exceeded 0.33, which means that, when comparing differentially corrected to raw measurements, even in the worst case the net improvement is about 70%. The Wilcoxon Matched-Pairs rank test clearly indicated that, in the open ($p < 0.05$) as well as in the forested area ($p < 0.05$), the results of the differential correction method are statistically better than those of the linear correction method.

IV.C.5 Discussion and conclusions

We clearly demonstrated the unreliability of the effect of the linear correction procedure on the accuracy of GPS measurements. A non-parametric multiple comparisons tests revealed a statistically significant difference in accuracy between linear and differential correction, with the residual positional error clearly reduced in the latter case. This can be explained by the difference in mathematical algorithms and data input procedures. To correct differentially, the base station calculates the expected time-lapse a signal needs to travel from a satellite to its receiving antenna. The difference between this expected travel time and the time actually measured (pseudo range) is the *time error*. This *time error* is computed at regular intervals (depending on the logging rate of the base station) for every 'visible' satellite and this information is stored. To adjust the raw measurements of a receiver (or rover) in the field, positional fixes are recomputed taking into account these time errors. Note that it is a prerequisite that all satellites used by the rover are equally 'visible' to the base station, so the computed time errors are valid for the respective field fixes. Linear correction, on the other hand, is based on the principle of calculating a ***positional error***. The rover, as well as the base station, uses the best combination of four available satellites (depending on the PDOP). Due to possible obstructions in the rover's line of sight, the satellite combination is not always the same as for the base station, resulting in inappropriate corrections often returning larger errors than those implicit in the original raw measurements. It is possible to control whether the same satellite constellation is used by both GPS receivers and to use only those measurements. Doing this, depending of the distance between the receivers and the operational conditions of the GPS in the field, one would have to omit a certain number of observations. Not once the same satellite constellation was used by both receivers on the location in the forested area. In the open field the number of observations using the same satellite constellation varied between 10 en 100%. This means that non of the observations in the forested area could have been linearly corrected when we would have imposed the condition that both GPS receivers should have used the same satellite constellation. For obvious reasons we decided against doing this and tested the accuracy (using the median) of linear correction performed on the total number of observations used in each data session because this mimics best the reality of using linear correction to correct animal borne GPS fixes.

Though we used the median of the locational errors as an approximate estimator to compare the accuracy of the different methods, one must not lose sight of the fact that, in wildlife research, not the median of multiple fixes but very often only a single fix is used to determine the animal's positions, resulting in even larger locational errors.

We therefore conclude that, because of the unpredictable outcome of the linear correction procedures, it is more appropriate to remain with raw GPS measurements, with the knowledge that 95% of the positions are expected to fall within 100 m of the true planimetric location.

IV.D Comments

Due to the fact that the Selective Availability was the accuracy of non-corrected GPS measurements improved significantly starting the first of May 2000. New tests on the difference in accuracy between non-corrected, linearly corrected and differentially corrected GPS measurements could be a topic for further research. The results of these accuracy-tests will also have a strong impact on the evaluation of GPS collars as an appropriate tool for studying roe deer habitat use.

The last part of this chapter (IV.C) was published as 'Casaer, J., Hermy, M., Verhagen, R. and Coppin, P. 1999: Appropriateness of the linear correction method for GPS positional fixes in wildlife studies. Wildlife Biology 5: 125-128'.

V Analysing space use patterns by Thiessen polygon and triangulated irregular network interpolation: a non-parametric method for processing telemetric animal fixes

V.A Abstract

This paper describes a new non-parametric model to estimate space use. We tested the similarity in space use patterns between an artificially generated utilisation distribution (Seaman - UD) and the model estimates for five different sample sizes. Three different test statistics (Kappa, γ, Tc) revealed a high similarity between the estimates and the Seaman - UD. The tests have also shown that, beyond a certain threshold value, larger sample sizes did not return significantly better results. We further demonstrated that the Thiessen model delineated areas of intensive use more effectively in a home range than Kernel estimators. UD estimation was done in AML (ArcInfo Macro Language), thus facilitating overlay operations with habitat information stored in the same ArcInfo GIS environment.

V.B Introduction

Analysing home ranges and habitat utilisation in wildlife studies usually involves three steps:

-Collecting the fixes (animal positions) in the field (trapping, radio telemetry, visual observation).

-Deriving the space use from those fixes.

-Overlaying space use patterns and habitat maps to analyse home range composition and habitat preference.

This chapter specifically deals with step two, the derivation of space use patterns from a set of fixes. This information can subsequently be used for home range analysis in habitat utilisation studies, as well as for other investigations requiring space use patterns, e.g. interactions between individual animals or species.

Several papers have been published reviewing methods and models to derive home ranges (Van Winkle 1975, Worton 1987, White and Garrot 1990, Harris et al. 1990, Kenward 1992). Home range models can be divided into parametric and non-parametric models, the latter being more robust than the former because they do not assume an underlying distribution (Worton 1987, Harris et al. 1990, Wray et al. 1992a). They can also be grouped into those defining the borders of the home range based on a percentage of the fixes to be included in the home range, i.e. Minimum Convex Polygon (White and Garrot 1990), Grid Cell Count (Voigt and Tinline 1980), Dirichlet Tesselations (Wray et al. 1992a) and Cluster Analysis (Kenward 1987), and those using the utilisation distribution or UD, a probability density function of locations to define the borders of the home range, i.e. Grid Cell Count (Voigt and Tinline 1980), Ellipses (Van Winkle 1975), Harmonic Mean (Dixon and Chapman 1980), Fourier Analysis (Anderson 1982) and Kernel Distributions (Worton 1989).

An analysis of the internal home range structure necessitates the application of methods that estimate the UD. It is critical that the UD be derived from the distribution of the fixes without any *a priori* assumptions about the nature of the UD. Therefore one cannot use models that assume normal or unimodal distributions of the animal fixes. Consequently only Kernel Distributions, Fourier Analysis, Grid Cell Counts, Cluster Analysis and Dirichlet Tesselations can be considered appropriate to study the internal home range structure.

The results of Kernel Distributions are not only highly dependent on ancillary parameters that must be specified by the user (Harris et al. 1990, Cresswell and Smith 1992, Wray et al. 1992b, Worton 1995), but also on the mathematical algorithms and grid sizes implemented in the software that computes the kernels (Lawson and Rodgers 1997). Fourier Analysis requires sophisticated mathematical computations (Harris et al. 1990). A common drawback of both methods is that all positional fixes must be independent. The outcome of the Grid Cell Count method, on the other hand, depends to a major extent on the size of the grid cells and the grid orientation (Worton 1989). The use of Dirichlet Tesselations was only recently proposed and, to our knowledge, has not been operationally applied, nor scientifically well developed.

We propose a new non-parametric method to estimate space use patterns based on Dirichlet Tesselations, which we further refer to as the Thiessen model (or HR95). The method does not make any assumptions with regard to the underlying distribution, and is directed at wildlife space use analysis where positional fixes are collected via automated radio telemetry systems or animal-borne GPS. Such data acquisition methods deliver highly correlated data points in large quantities (see for example Rempel et al. 1995 and chapter IV). This Thiessen model has been validated against a Seaman-generated space use distribution (Seaman 1996) and cross-referenced against Kernel estimators.

V.C Model and Tests

V.C.1 Model

V.C.1.a Subdividing the study area

The Thiessen model starts by subdividing the study area in Thiessen or Voronoi polygons, also known as tiles. One such tile comprises that segment of the total area that encompasses all locations closer to a particular animal position or fix than to any other fix (Wray et al. 1992a). In other words, a tile defines or delineates subzones within an animal's home range area that are exclusively assigned to a particular fix. The mathematical expression of the Thiessen polygon is the following:

$$T_1 = \left[x : d(x, P_1) < d(x, P_m); \forall m \neq 1 \right]$$

where: x = a 'location' in the research area

P_1 = position 1 (in this case a fix)

T_1 = polygon or tile for position 1

P_m = any other position

d = Euclidean distance.

Whenever several fixes are close to each other, an aggregation of small Thiessen polygons is generated. In contrast, isolated fixes result in larger Thiessen polygons, as is illustrated in Figure V.1.

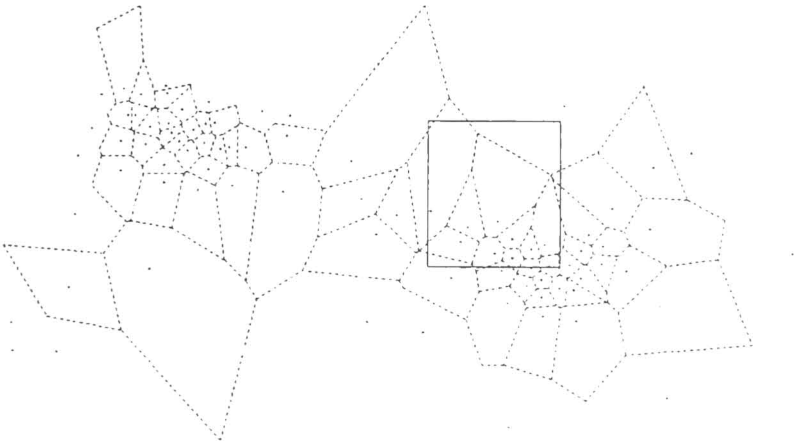

Figure V.1: Division of an area in Thiessen polygons, showing an aggregation of small tiles where several positional fixes occur together. The black dots represent animal fixes. The box identifies the area enlarged in Figure V.2.

V.C.1.b Computing the space use intensity

The inverse of the area of a Thiessen polygon (area $^{-1}$) can be considered a measure of terrain utilisation, referring to the number of fixes per area unit, and further called intensity. This measure is comparable with the density as computed in the Grid Cell Count methods. Intensity values are then assigned to the polygon-specific fixes.

Note that the tiles around the outermost fixes are constrained by the overall predefined Thiessen interpolation zone margins. These in turn are derived from the geographical locations of the outermost fixes and a default value for margin size (ESRI 1992).

Without modification this would result in intensities that have no real meaning nor correlation with the real number of fixes per area unit. To avoid this, an intensity value of zero is assigned to the outermost fixes, thus automatically and empirically defining an external home range border. As this external border includes all study area fixes or observations, it can be considered the equivalent of a home range delineation as prescribed by the Minimum Convex Polygon method.

V.C.1.c Generating a 3D surface and space use patterns

Subsequently, a triangulated irregular network surface or TIN (Van Kreveld 1996) is constructed out of the irregularly spaced intensity points (Figure V.2).

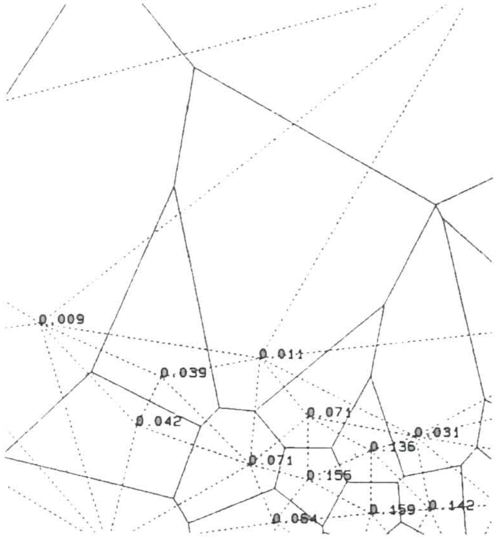

Figure V.2: Thiessen polygons and TIN interpolation based on use intensity values assigned to the original fixes. The dots locate animal fixes and the use intensity values represent area^{-1} of the Thiessen polygons (area enlarged from Figure V.1).

The TIN then effectively serves as a three-dimensional space use model where *x* and *y* define spatial characteristics (locational constraints) and z the thematic attribute (intensity of space use). See Figure V.3 for a visual impression.

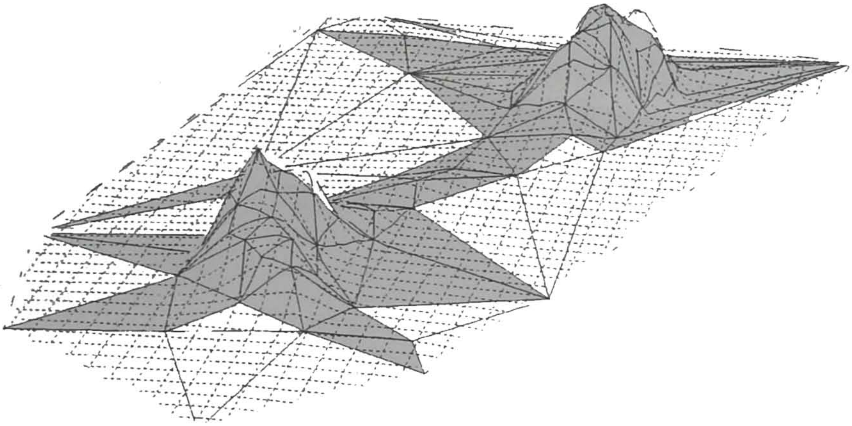

Figure V.3: Three-dimensional model surface for habitat use (area identical to that of Figure V.1)

The 3-D model is used to estimate the structure of the 'home range', which is defined as the minimal area including 95 or 90 percent of the total volume (Anderson 1982, Boulanger and White 1990, White and Garrot 1990). Internal space use areas can be calculated in an analogue manner, representing the areas delineating 85, 80, ...,15,10 and 5 percent of the total volume (Figure V.4).

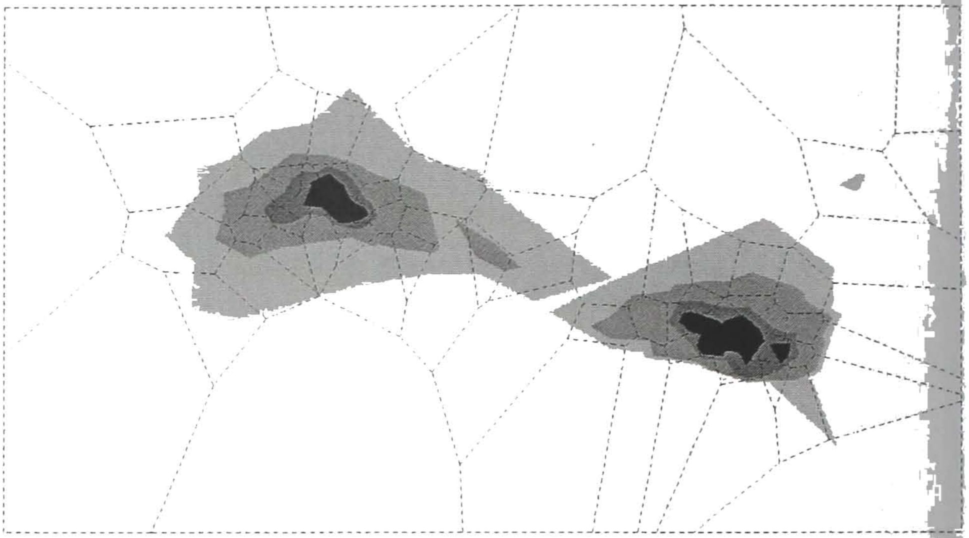

Figure V.4: UD represented as a function of the volume of the total UD (the contours represent the different areas including a percentage of the home range volume, respectively 95%, 75%, 50% and 25%).

The program code to analyse space use from the distribution of the animal fixes was written in AML (ArcInfo Macro Language), facilitating a direct overlay with habitat information in the ArcInfo vector GIS environment. The program code is directly compatible with ArcInfo version 7.1.2 (ESRI, 1997) for Windows NT or UNIX, and is available upon request from the author.

V.C.2 Tests

To evaluate the method's ability to estimate an UD, the true utilisation distribution, or a very good model thereof, has to be known. Since the true UD cannot ever be extrapolated from a sample of animal fixes, not even from an extensive one, we had to opt for artificially generated UD's as reference material. To that extent we made use of a UD generating program that produces composite bivariate normal utilisation distributions (Seaman 1996). Complex UD simulations such as Seaman's often have several activity nuclei (Kenward 1992). They are therefore considered to quite adequately mimic space use, especially where it concerns mammals that exhibit patchy habitat use (Seaman 1996, Lawson and Rodgers 1997). Seaman's program determines at random how many bivariate normal distributions will be put together and to which extent they will overlap, allowing the user to select the number of fixes, sample size being the variable. Based on a certain UD, the program can produce as many replicates as desired for a chosen number of fixes. For our study we made the Seaman program generate nine independent replicates for each of five sample sizes (200, 400, 600, 800 and 1000 fixes), returning a total of 45 UD estimates. To analyse the similarity between the Seaman generated UD and the equivalent Thiessen models, Kappa coefficients of agreement, γ-statistics and Kendall Tc's were computed.

Kappa coefficients of agreement are numeric indicators for the similarity between two grid-layers, excluding the agreement that occurs purely by chance.

$$K = \frac{P(A) - P(E)}{1 - P(E)}$$

P(A) depicts the proportional agreement between grid cells, while P(E) is a proportional measure of the agreement by chance only (Siegel and Castellan 1988). Complete agreement between a UD estimated by the Thiessen model and the equivalent Seaman UD results in a Kappa value of 1. Because Kappa coefficients of agreement are raster-based, the respective UD pairs had to be reformatted into 600 x 800 cell grid layers, and reclassified into five space use intensity categories. Kappa's were then computed within the analysis module of IDRISI (IDRISI 1997) for each of the 45 UD pairs (Seaman's + Thiessen's).

Though the Kappa statistic is a good measure of agreement between data sets with identical locational characteristics (e.g., grid layers), there are some caveats. Enlarging the search radius within which the grid layers are cross-referenced results in a larger number of grid cells not belonging to the areas of intensive use (usually located centrally within the model), both for Seaman and Thiessen UD's. As such, the number of grid cells of identical value in the Seaman as well as in the Thiessen UD is artificially increased giving larger kappa coefficients of agreement. Kappa is also highly dependent on the number of classes cross-referenced. Reducing the number of classes decreases the probability of unequal cross-referencing or wrong assignment thus potentially increasing Kappa (Coppin and Bauer 1994).

As alternative measures of agreement between the respective UD's, γ and Kendall T_c statistics were derived for the original vector data in an ArcInfo environment. Intensity values were sampled gridwise (25 x 25 or 625 values) as to cover the full geographical extent of the UD's to be cross-referenced, and categorised into 21 classes (zero class and twenty 5% classes). γ Statistics comparing the Thiessen and Seaman values were computed according to the formula proposed by Goodman and Kruskal (1954):

$$\gamma = \frac{\Pi_s - \Pi_d}{1 - \Pi_t}$$

where: $\Pi_s = \Pr\{a_1 < a_2 \text{ and } b_1 < b_2; \text{ or } a_1 > a_2 \text{ and } b_1 > b_2\}$

$\Pi_d = \Pr\{a_1 < a_2 \text{ and } b_1 > b_2; \text{ or } a_1 > a_2 \text{ and } b_1 < b_2\}$

$\Pi_t = \Pr\{a_1 = a_2 \text{ or } b_1 = b_2\}$

a_i = the estimated value

b_i = the reference value

The γ statistic characterises 'the difference in the conditional (no ties) probabilities of like and unlike' (Siegel and Castellan 1988). It is a more appropriate parameter to measure association between ordered variables when many ties occur than, for example, the Kendall T_b or Spearman statistics.

In addition we applied the Kendall T_c test as defined by Stuart (1953):

$$Tc = \frac{\Pi_s - \Pi_d}{(m-1)/m}$$

Here m is the minimum of α and β, whereby α and β represent respectively the number of rows and columns in the contingency table, and Π_s and Π_d are defined as in equation 3. Kendall's Tc is a correction for Kendall's T (Kendall 1938) whenever many ties occur a contingency table (Stuart 1953).

Finally a Kruskal-Wallis test and subsequent multiple comparison tests were implemented to compare the value pairs obtained for the different sample sizes for each of the three test statistics (Siegel and Castellan 1988).

V.C.3 Comparing the Thiessen model and Kernel estimators

To compare our Thiessen model to the Kernel estimator, the latter apparently being the more widely used method (Worton 1989, Worton 1995, Seaman 1996, Lawson and Rodgers 1997), we generated four simulation data sets from sequential series of fixes. Two were based on 1000 fixes and made to agree with the distribution patterns observed during fieldwork in the summer of 1997. The first thereof mimicked the preference for forest edges and corridors (Figure V.5a), the second the avoidance of a large open field. The two other simulations were based on 6000 fixes and generated via a "random walk" algorithm (example in Figure V.5b).

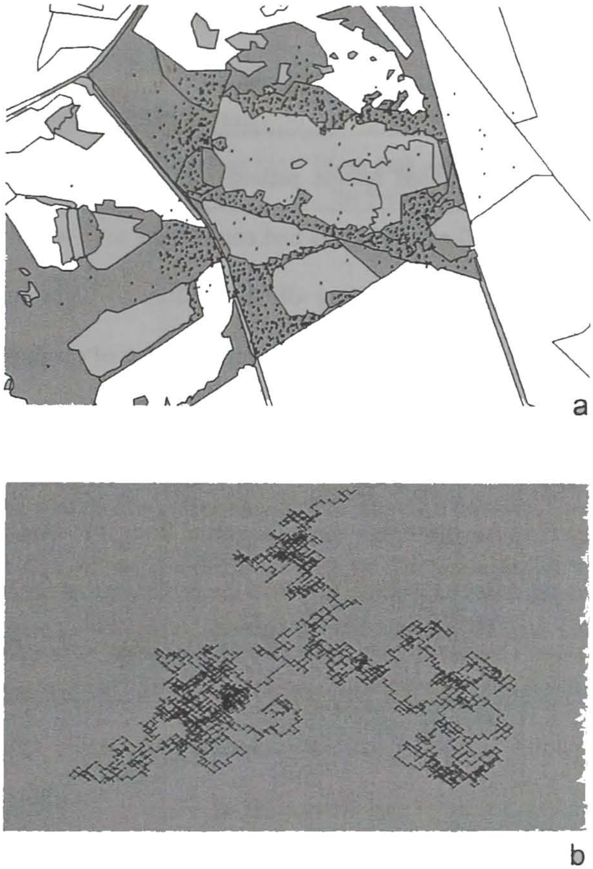

Figure V.5: Graphical representations of two examples of the data sets used to compare the performance of the Thiessen model against that of the Kernel Estimator. The first figure (a) depicts a set of 1000 fixes based on the preference criteria, the second (b) a set of 6000 fixes generated by a random walk algorithm.

Uniform samples of 10% (and in addition also 5% for the random walk simulations) of the total number of fixes were extracted by choosing each tenth (and twentieth) fix, this to duplicate the automatic sampling procedure typical for GPS and automated radio telemetric systems. The sampling resulted in data sets containing or 100 fixes (first two simulations), or 300 and 600 fixes ("random walk" simulations). Because

Schoener index values were significantly below 2, autocorrelation as defined by Swihart and Slade (1985) could be considered high or outspoken.

To investigate which model performed best in delineating areas of high intensity use, we overlaid the sample-based isoline patterns of space use on the corresponding original datasets of 1000 or 6000 fixes and counted the total number of observations enclosed by single isolines (including all observations of higher use intensity when present). We then computed the ratio of the total number of observations enclosed by an isoline over the total area delineated by the same isoline and used the change in this ratio as a measure to compare the Thiessen model and the Kernel estimator. This was necessary because the home range program RangesV (Kenward and Hodder 1996) we used generates its isolines based on the percentage of observations included, in contrast to our Thiessen model that generates isolines in function of the volume included.

V.D Results

V.D.1 Kappa coefficients, γ and Kendall T_c test statistics

The three test statistics used to measure the agreement between the Seaman generated and the Thiessen estimated space use patterns returned very similar results. Nevertheless, a Kruskal Wallis test revealed considerable differences for the different sample sizes among the individual Kappa's, γ's and T_c's. A multiple comparisons test, where all the mean rank values were cross-referenced simultaneously for each of the test statistics, allowed us to determine which differences were indeed statistically significant. An overview of the mean rank values (used in the multiple comparisons tests) for the different sample sizes and test statistics is given in Table V.1.

Table V.1: Mean rank values for the different sample sizes for each of the test statistics, computed from nine replicates

Number of fixes in the samples	Mean rank values for the Kappa's	Mean rank values for the γ–coefficients	Mean rank values for the T_c-coefficents
200	6.56	7.67	5.67
400	14.89	24.22	18.44
600	24.89	23.11	23.78
800	33.89	30.56	32.56
1000	34.78	29.44	34.56

The threshold for statistical significance of the difference between a pair of mean rank values of a test statistic (at $\alpha = 0.05$) was 17.38. The table shows that, while there existed no significant differences between sample sizes of 200 and 400, 400 and 600, 600 and 800, and 800 and 1000 fixes, sample size increases from 200 to 600 or more fixes always (with a very slight exception for the γ statistic) resulted in significant differences. Thiessen UD's based on the sample size returning the higher mean rank value match the patterns of the equivalent Seaman UD's significantly better than those with lower mean rank values.

V.D.2 Thiessen model versus Kernel estimator

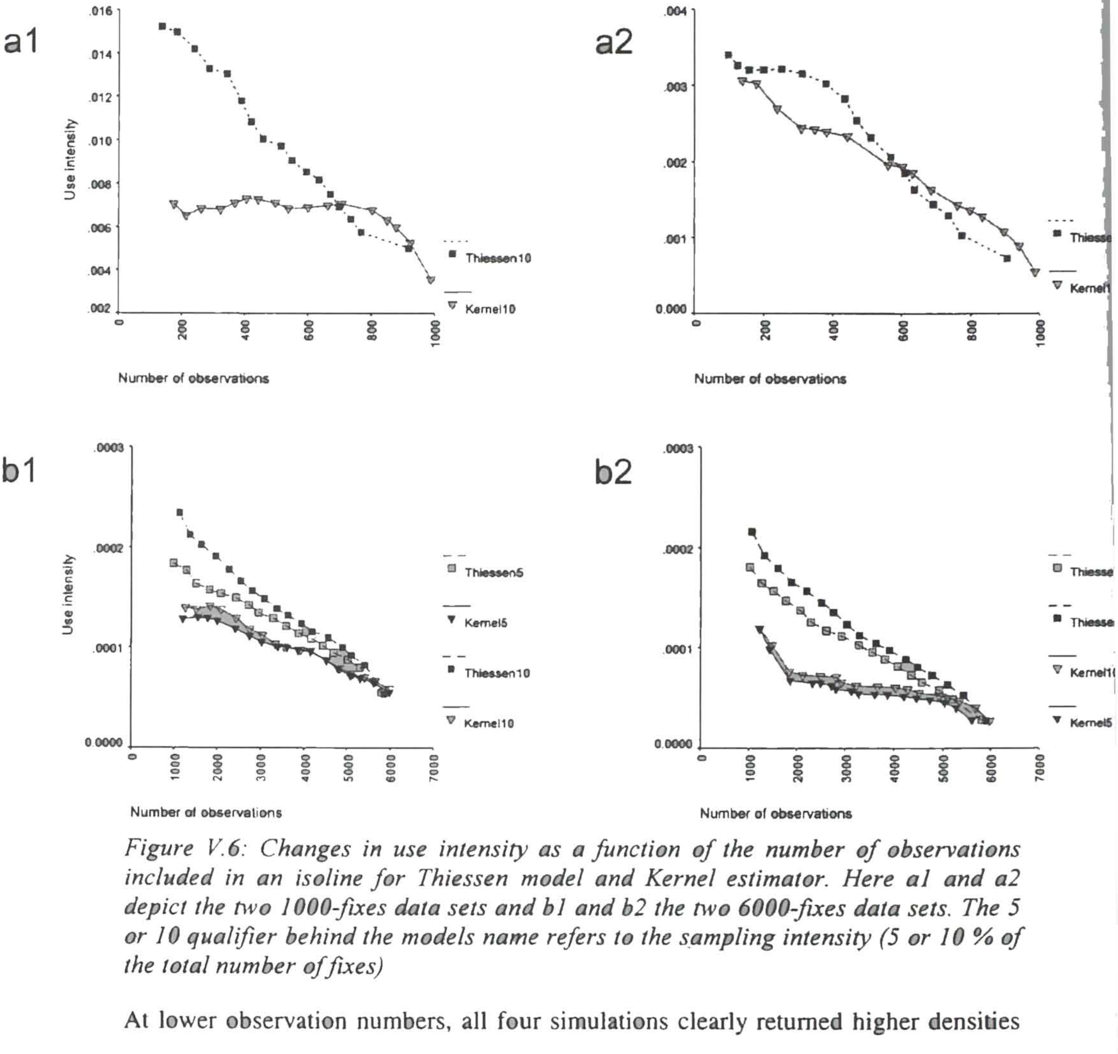

Figure V.6: Changes in use intensity as a function of the number of observations included in an isoline for Thiessen model and Kernel estimator. Here a1 and a2 depict the two 1000-fixes data sets and b1 and b2 the two 6000-fixes data sets. The 5 or 10 qualifier behind the models name refers to the sampling intensity (5 or 10 % of the total number of fixes)

At lower observation numbers, all four simulations clearly returned higher densities for the isolines delineated by the Thiessen model than for those delineated by the Kernel estimator (Figure V.6), reflecting the fact that the Thiessen model did a better job in identifying the areas of intensive use or core areas. The Kernel estimator, on the other hand, appeared to have included larger chunks of less-intensively used areas in order to attain the overall percentage of observations required to demarcate isolines of equal use. It is also evident that the decrease in density, with increasing numbers of fixes, is more outspoken for the Thiessen model than for the Kernel estimator, especially at the lower numbers of fixes. This highlights the fact that the Thiessen-delineated core areas have a relative higher use intensity than the Kernel-estimated core areas, when compared with their respective other usage zones.

V.E Discussion and conclusions

The literature of the last decade encompasses several manuscripts covering the capabilities of different methods to portray internal home range structure and/or to delineate core areas, or other aspects based on internal space use patterns (Worton 1987, Harris 1990, Kenward 1992 and Wray et al. 1992b). But one also learns from that same literature that quite often the first, and not seldom the sole characteristic that is compared when cross-referencing different home range models, is the total home range size and/or shape (Boulanger and White 1990, Worton 1995, Lawson and Rodgers 1997).

While a comparison of different home range models may be of interest, the existing discrepancies in home range estimation methods greatly reduce the value of such an analysis. In a recent article, Lawson and Rodgers (1997) discussed the large differences in home range sizes that can occur, even when using the same 'method' to calculate the home range. These differences may be caused by the implementation of different software packages or, more importantly, by the use of different parameters to compute the home range size. To adequately evaluate the ability of a method or program to estimate space use patterns, more information is needed than only home range size and/or shape.

We propose that the analysis of the similarity between the isoline patterns generated by a theoretical UD (e.g., Seaman simulation) and a model estimate (e.g., Thiessen model) represents a valuable technique to compare and cross-reference different home range models. Moreover, our results clearly give evidence of the fact that the Thiessen model presented here, is a good estimator of a Seaman's simulation of space use in case a true UD is unknown or unavailable.

Neither Kappa, nor the γ or T_c test statistics benefited significantly from sample size increases beyond 600 fixes. This does, however, not imply that the Thiessen model we have presented here always requires a minimum number of 600 fixes to adequately estimate an unknown UD, nor that 600 fixes is always the threshold value above which no further changes in the results will be found. The number of observations required to estimate a UD not only depends on the method, but also to a major extent on the true, generally unknown utilisation distribution. Estimating a unimodal, bivariate normal distribution evidently requires fewer fixes than estimating a multimodal, non-normal distribution. Because of their nature, non-parametric models often need more data points to estimate complex UD's (Worton 1987, Kenward 1992). We may therefore conclude that the determination of the exact number of positional fixes required for estimating space use intensity is a rather unrealistic goal unless knowledge on the true UD is available. However, the intensified use of automated radio tracking systems in the near future will undoubtedly force us to consider the trade-off between the frequency of fixes, and the length of the total observation period, rather than the total number of fixes required to solve a research question. If this trade-off is not investigated properly, enormous data sets will be made available to the researcher, probably with little to no additional benefit.

The comparison of the Thiessen model to the more traditionally used Kernel estimator allowed us to formulate an answer to the question as to which of the two approaches better detected and delineated areas of intensive space use. We based the comparison on samples extracted from an underlying utilisation distribution. Sample selection mimicked "natural" sampling procedures incorporation concepts of preference, avoidance, and random-walk, and sampling intensity was chosen to be realistic from the perspective of future home range data collection with high technology tools. Our results point to the Thiessen model as being more appropriate than the Kernel estimator to spatially differentiate space use intensity in the underlying UD via isolines of equal space.

Though for this comparative study sample sizes largely exceeded those of many traditional radio telemetric studies (exceptions occurred, e.g., Koubek 1995), the forthcoming use of animal-borne GPS instruments will result in even larger sets of highly correlated fixes. The development of methods not making any *a priori* assumptions about the distribution of the unknown UD while being insensitive to the auto-correlation among animal fixes, is a prerequisite for the analysis of such data sets. We designed this approach as a first phase in this process. This may furthermore not only lead to a better understanding of the actual patterns of intensive space use, but will also definitely and more immediately bring about an improvement in UD estimation approaches.

These results were published as ' Casaer, J., Hermy, M., Coppin, P., and Verhagen, R. 1999. Analysing space use patterns by Thiessen polygon and triangulated irregular network interpolation: a non-parametric method for processing animal fixes. International Journal for Geographic Information Science. 13 (5): 499-511 '.

VI Analysing habitat use by roe deer fawns (*Capreolus capreolus*) using different criteria

VI.A Abstract

The use of edges by animals has been the topic of wildlife research for many years. In this paper the use of transition zones between different habitat types by roe deer fawns was studied.

For this purpose eight roe deer fawns were radio tracked daily during the first months of life. The highly auto-correlated radio fixes were analysed using a self-developed home range program (HR95) and minimum convex polygons.

Compositional data analysis was applied to analyse the preference or avoidance of the animals for transition zones. Furthermore, the edge density and the density of patches in the home ranges of the fawns were analysed.

There was a clear avoidance of areas that were further than 45 meters away from any edge. The fawns were using home ranges characterised by a significantly higher edge- and patch-density than other parts of the study area.

We showed the great influence of the definitions applied to determine the 'used and the available habitats' on the outcome of compositional data analysis, and showed the importance of using biologically correct and fine-tuned home range programs.

VI.B Introduction

Habitat use by roe deer has been studied throughout Europe over the last 20 years using radio telemetry (*Denmark:* Jeppesen 1990, *France:* Bideau et al. 1987, Bideau et al. 1993, Maillard et al. 1999, *Norway:* Aanes and Andersen 1996, Mysterud 1998, Tufto et al. 1996, *Sweden:* Guillet et al. 1996). However, few studies exist at present on habitat selection by roe deer fawns, or on the factors affecting their habitat selection.

The shy behaviour of roe deer fawns during their first weeks makes it very difficult to study them. Roe deer is classified as a 'hider' species (Geist 1981, Lent 1974). The fawns rely on their perfect hiding capacities to survive the first weeks of life. Fawns "freeze" when approached by predators. Only as they become older and quicker they try to escape when disturbed. According to Linnell it is only at 40 days old that the majority of fawns flush when a person approaches within 10m (Linnell et al. 1998b). Other published work on roe deer fawns concerns predation (Linnell et al. 1995, Aanes and Andersen 1996), home ranges size, daily movements and behaviour (Boutin et al. 1991, Van Laere et al. 1996) or the characteristics of bedsites (Gaillard and Delorme 1989). According to several authors, the choice of bedding sites is a key element in the anti-predator strategy during the first weeks of life (Mysterud and Ostbye 1999). Other authors suggest that the selection of the bedding sites is, in contrast, linked with thermoregulation (Gaillard and Delorme 1989).

Only Linnell et al. (in press) compared the habitats where bedsites were found with the habitats available in the study area (second order selection (Johnson 1980)), and within the individual home ranges (third order selection (Johnson 1980). None of the above-mentioned studies focuses on the role of edges in the habitat selection by roe deer fawns, though the importance of edges for adult roe deer has been proven both without telemetry (Reimoser and Gossow 1996) and by radio telemetric studies (Tufto et al. 1996). Our research aims to analyse the preference of roe fawns for edges both on the second and third order of selection.

During the last decade, many studies were published on the differences between various home range methods that can be used to study habitat selection (Boulanger and White 1990, White and Garrot 1990, Lawson and Rodgers 1997, Casaer et al. 1999a). Most of them focus on the theoretical background, explaining why certain models are more valid than others to derive the patterns of space use from animal locations collected by radio telemetry.

Other papers compared the different methods for the analysis of habitat selection (Alldredge and Ratti 1986 and 1992, Manly et al. 1993). They point out which statistical pitfalls should be avoided when analysing home-range composition and animal locations to study habitat preferences. Nevertheless, few biological studies take this theoretical information into consideration when analysing real data.

We therefore compared the results of using different home range models as input for the compositional data analysis of the locations of the eight roe deer fawns. GIS analysis tools (Fragstats) were used to compare the fawns' home ranges with other areas in the study area. The outcome of the GIS analysis was also compared with the results of the compositional data analysis. The overall aim of the research was to analyse if roe deer fawns show a preference for habitat transition zones and to study the sensitivity of compositional data analysis to the methods used as definitions for 'use and availability'.

VI.C Study Area

The study area was a military exercise zone (Wuustwezel – Brecht) in the north of Flanders (Province of Antwerp, Belgium). The total area covers approximately 1500 ha, and is surrounded by small, private forests. Of the study area, 45% consists of open habitats such as heathland, grassland, or a mixture of both vegetation types and 35% is forested, mostly pine (*Pinus sylvestris* or *Pinus nigra*) stands or otherwise a mixture of oak (*Quercus robur*) and birch (*Betula spp.*). The forest stands are characterised by much variation in undergrowth. The whole study area is closed to the public. Hunting is forbidden in the central part of the study area but allowed in the surrounding areas. The mosaic of natural and semi-natural habitats surrounded by cultivated land, meadows and human settlements is typical for the actual situation of most nature areas in Flanders.

VI.D Methods

VI.D.1 Habitat mapping, edges and cover

Based on digital aerial photographs, topographical maps and the existing maps representing the different forest stands a GIS layer was generated containing the different habitat patches in the study area. The boundaries of the habitat patches, as well as the assigned habitat types, were verified in the field and, when necessary, corrected using GPS–mapping methods. We distinguished 21 habitat types, including five different forest types (Table VI.1).

We measured the cover in the different habitat types using an adapted version of the cover pole method described by Griffith and Youtie (1988). Using smaller measuring intervals and using the potentially visible proportion of a roe deer in each height interval (see chapter III), we achieved a good estimate of the hiding cover.

To study the edge effect we first generated a buffer of 15 meters on both sides of each transition between two habitat types. Subsequently, a buffer of 30 meters and one of 45 meters were generated. Thus, the whole study area was subdivided into four classes of buffers; areas that are more than 45 meters away from any transition zone, those that are between 45 and 30 meters, between 30 and 15 and less than 15 meters respectively away from any edge (Figure VI.1). Originally the intention was to create a second buffer map, buffering only the edges between habitats with different cover values. Indeed, Reimoser showed the attractiveness for roe deer of clear visual edges between patches providing dense cover and those that do not (Reimoser and Gossow 1996). However, due to the large variation in measured cover values within a single habitat type (see also chapter III), it was impossible to translate the habitat map into a cover map, which would have been necessary to generate this second buffer map.

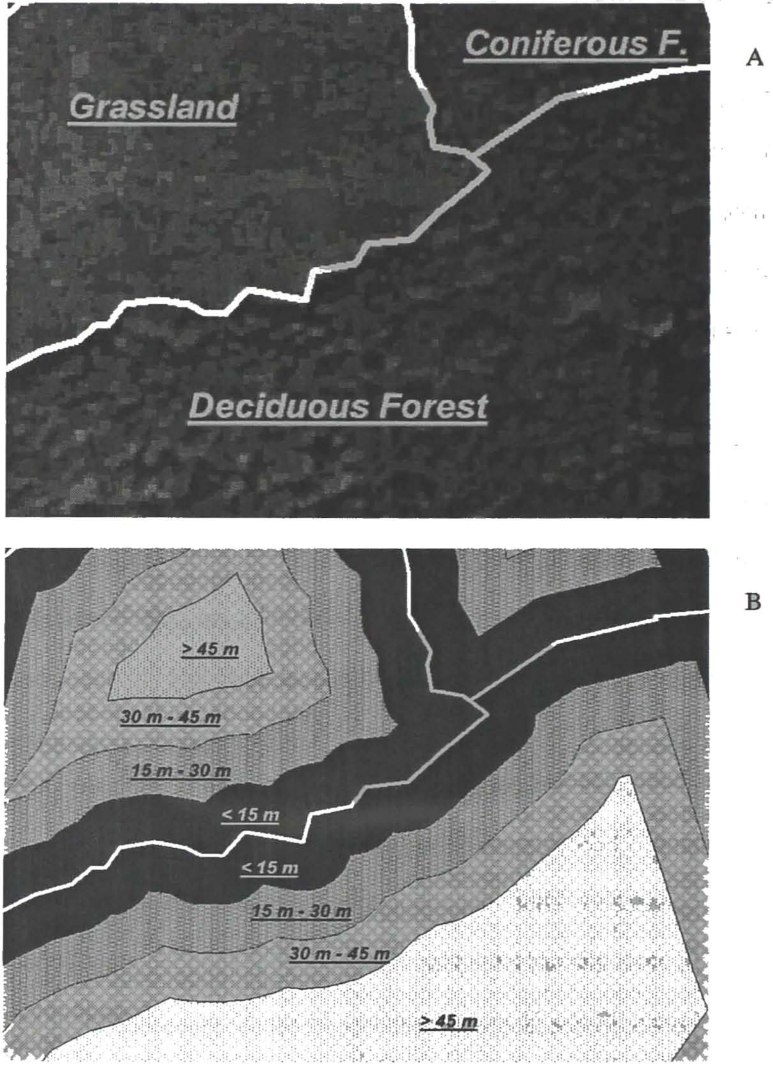

Figure VI.1: An example of an aerial photograph and delineated habitat patches (a) and the derived buffer map (b).

VI.D.2 Radio telemetry

During the months of May and June 1997 and 1998 we searched for roe deer fawns using daily car drives in order to locate pregnant does. Once these were located, we tried to observe the does each day until the fawn was discovered. We combined this method with the use of roe calls as described by Van Laere et al. (1989). Though this method sometimes gave us a direction to start the search for fawns, most of the fawns were found by random searches in the direct vicinity of the places were the does were observed to behave particularly nervously. Once the fawns were found we equipped them with expandable radio collars (TXH-2, Televilt). The lifetime of the collars is expected to be 15 months (nevertheless at least one of the collars was still working in the summer of 2000, more than 24 months after being started). Those animals judged to be too small to be collared were marked using ear tags. During the second year, the fawns that were considered to be too small to be collared, were marked using ear tags with a small rat transmitter glued on. This enabled us to relocate the animals in subsequent days and then to fit them with an expandable radio collar.

The radio-collared fawns were located every day using an RX 8910 or an RX 900 receiver with respectively a built-in H antenna and a Yagi antenna (Televilt). We did not use triangulation since an error of less than 10 meters was required. Visual confirmation showed that our tracking method allowed the location of bedding sites or even active fawns with an accuracy better than 5 meters.

To determine the geographic co-ordinates of the animal locations, we used a Trimble ProXL GPS–rover, in combination with a base station. We averaged 30 differentially corrected positions, with a logging interval of 5 seconds, to achieve an accuracy better than 1 meter (Sigrist et al. 1996). The location error on the animal fixes is therefore due to the radio telemetric procedure and can be said to be less than 10 meters.

In contrast to other studies (Van Laere et al. 1996) we continued to locate the animals every day until they were approximately three and a half months old. Fixes collected afterwards were not included in the analysis as the fawns change their behaviour. Linnell described that towards this age, roe deer fawns become more and more active. After the summer rut all traces of the hiding behaviour fade and roe deer fawns are most of the time associated with the doe (Linnell 1994, Linnell et al. 1998b).

VI.D.3 Space use

Recently De Solla et al. (1999) stressed the biological importance and relevance of using all the locations, rather than eliminating observations to achieve independent observations, which is required by most of the commonly used home-range programs. Most of our observations sets were highly auto-correlated. We therefore analysed the space use patterns of the roe deer fawns by applying only methods that are not influenced by the auto-correlation between daily fixes (see chapter V).

Firstly, we used the regularly used Minimum Convex Polygon (MCP) and the Minimum Convex Polygon based on 95 percent of the locations (after removing the 5% most extreme outliers) (MCP95). This enabled direct comparison of our results with other studies. To calculate the MCP and MCP95 we used the program 'MOVEMENT' an extension module for ArcView (Hooge and Eichenlaub 1997 v.1.1.).

Secondly we used a space-use program that was developed by the author (Chapter V). The program delineates the home range (HR95) as the area including 95% of the volume of the estimated space use function (Casaer et al. 1999a). The core area (CA) is defined as that part of HR95 where the intensity of use is higher than the average use-intensity in HR95. The main advantage of the program is that it estimates a space use pattern based on the distribution of the animal fixes, without any *a priori* assumption on the nature of the underlying, unknown space use distribution. In contrast to the two other models it excludes large unused areas.

Though most fawns were radio collared between the 5th June and the 15th June, data were analysed in both years only for the period when there were daily locations for all four animals. In 1997 this was from the 6th July until the 21st September. In 1998 the period extended from the 30th June to the 21st September. Before and after these periods the animals were radio tracked intermittently. Since the home range model assumes that the locations are done at regular, constant time intervals, we omitted these data from the space use analysis.

VI.D.4 Home range characteristics

The fragmentation of the study area and of the roe deer home ranges was analysed by applying FRAGSTATS*ARC V2.0.3 (Innovative Inc.). These analyses returned the total edge length (TE), the total area (TA) and the number of patches (NP) as absolute measures. To compare the home ranges of the roe deer fawns with the overall study area, these measures were transformed to edge density (ED) and patch density (PD). These are the absolute measures divided by the area. Since comparing the characteristics of the eight home ranges with those of the total study area would, because of the small sample size, be statistically meaningless, we generated 60 random home ranges. The simulated home ranges were created in Arcview by generating 60 randomly distributed circles having an area distribution similar to the one of the eight home ranges. We a Wilcoxon-Mann-Whitney test, as well as a randomisation test (written in SPLUS 2000) to compare the characteristics of the true home ranges and the simulated ones. For the randomisation test we generated 1000 random sets of eight home ranges, drawn randomly from the total set of the simulated ones and the eight roe deer home ranges together. We compared the mean of the ED, the PD, the TE, the NP and TA of these 1000 sets with the mean of the ED, PD, TE, NP and TA of the set of eight roe deer home ranges. Because of the non-normal distribution of several of the characteristics we calculated the p-values and their confidence intervals as described by Hugueny and Guégan (1998).

VI.D.5 Habitat selection

We studied the habitat selection on the second and third order of selection (Johnson 1980). To study the second order selection we applied a design II; this implies that habitat availability was estimated for the whole population while the use of the different habitats was estimated for each animal individually. To analyse the habitat selection within the home range (third order) a design III was used. Here both use and availability were estimated for each individual animal (Manly et al. 1993). In both these set-ups the animal and not the individual observation is the sample unit.

To assess second order availability we used (*i*) the arbitrarily defined study area, (*ii*) the MCP based on the pooled locations of all the animals (MCPtot) and (*iii*) the sum of the MCP's of each individual animal (Sommcp). We defined the used area as (*i*) the MCP, (*ii*) the MCP95, (*iii*) the home range (HR95) and (*iv*) the core area (CA).

To analyse the third order selection the availability of the different habitat features was calculated in (*i*) the MCP, (*ii*) the MCP95 and (*iii*) the HR95. Use was defined firstly as the habitat features present in (*i*) the core area and secondly as the characteristics of the habitat for each of the animal fixes in (*ii*) the MCP and (*iii*) in the HR95.

Using the above-mentioned criteria to define use and availability, compositional data analysis was used (Aebischer et al. 1993) to analyse the habitat selection. This method deals with most of the problems and constraints, such as the unit-sum-constraint of the different proportions of use and availability, which limits the validity of other methods (Tufto et al. 1996). The compositional analysis was performed in a program called 'Resource selection' (Leban 1998). The program first tests the H_0 hypothesis that habitat use is proportional to habitat availability. When the H_0 is rejected at a significant level (0.1 or 0.05), the ranking matrix of the relative preference of the different habitat types is generated and the significance of the ranking is computed. When the available habitat types are disproportionally used by the animals, but with large individual differences in the selection, this results in a rejection of the H_0 but not in a significant ranking of the preference for the habitat types.

VI.E Results

VI.E.1 Habitat characterisation

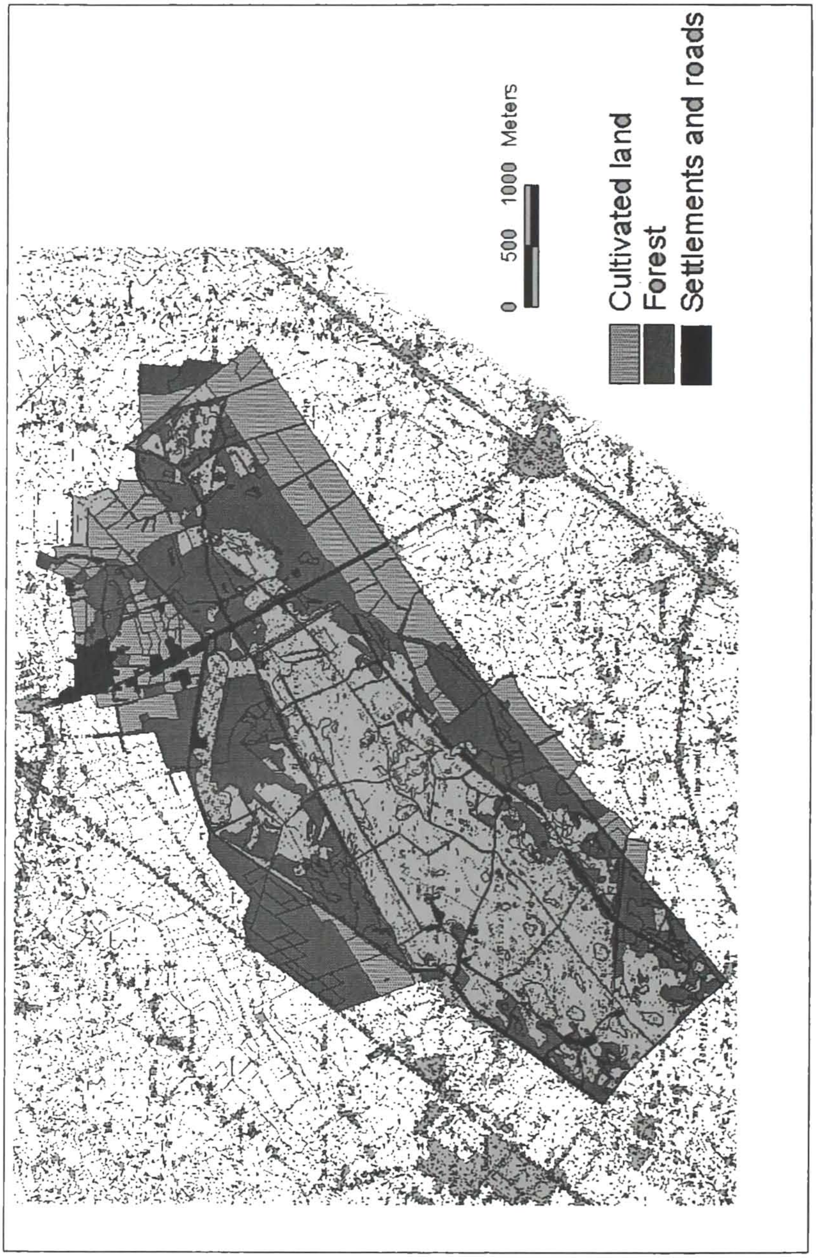

Figure VI.2: Simplified habitat map of the study area.

Table VI.1: The 21 habitat types used to describe the study area.

Habitat type	Area (ha)	Percentage of total study area
Heathland	10.19	0.7
Grassland	21.80	1.4
Mixture of grasses and heather	508.77	33.1
Tall grassland	2.25	0.1
Dense shrub vegetation	2.28	0.1
Open Shrub vegetation	35.52	3.3
Linear woody habitat-elements, along roads	16.72	1.1
Linear woody habitat-elements, between two natural habitat types	18.80	1.2
Meadows within forests	7.59	0.5
Feeding grounds	0.78	0.1
Cultivated land	31.96	2.1
Used meadows	239.44	15.6
Open areas in forest stands	19.33	1.3
Deciduous forest stands	227.59	14.8
Mixed forest stands	48.35	3.2
Coniferous forest stand	163.11	10.6
Trees and grassland	46.49	3.0
Others	24.04	1.6
Sandy open areas	1.64	0.1
Water	59.89	3.9
Swamp Vegetation	0.57	0.0
Roads	66.70	4.3

VI.E.2 Radio telemetry

In total 28 roe deer fawns were found and 19 of them were fitted with expandable radio collars. Though the exact day of birth was not known, we estimated, based on their behaviour and activity (see Linnell et al. in press), most of the roe deer fawns to be in their second week when they were fitted with radio collars. Over the two years, 11 (> 50%) of the fawns that were radio-tracked died before the 10th of July. Most fawns were found dead between the sixth and the twelfth day of radio tracking. Six of those 11 animals showed evidence of predation by fox (*Vulpes vulpes*), or wandering dogs. Two fawns died of malnutrition or cold weather at a very young age. Our results coincide with the results of Aanes et al. (1998), who found high predation rates between the ages of two weeks up till the sixth week after birth. Similar to our results, Linnell (1994) concluded that predation, followed by starvation/hypothermia were the major causes of neonatal mortality of roe deer.

VI.E.3 Space use

We monitored four fawns in 1997 for 77 days and four fawns in 1998 for 84 days. Table VI.2 gives the characteristics of the resulting MCP, MCP95, HR95 and core areas (CA) of the eight roe deer fawns.

Table VI.2: Sizes of home ranges and the numbers of animal observations within the home ranges.

Freq.	151.175	151.065	151.135	151.154	151.265	151.344	151.325	151.304	Average	StDev.	Conf. Interv.	Cv.
MCP (ha)	11	11	32	19	24	48	44	75	33.0	21.9	15.2	66.4
MCP95 (ha)	9	10	29	16	18	31	39	69	27.6	19.8	13.7	71.6
% area	75	85	91	84	75	65	89	92	81.9	9.6	6.6	11.7
HR95 (ha)	5	6	19	12	14	27	18	50	18.9	14.5	10.0	76.6
%area	47	52	60	63	58	56	41	67	55.5	8.5	5.9	15.3
%obs	79	72	72	77	82	74	71	73	75.0	3.9	2.7	5.2
Core Area (ha)	1	1	4	2	4	8	6	12	4.8	3.8	2.6	80.2
%area	11	11	12	11	17	17	14	16	13.5	2.6	1.8	19.3
%obs	56	48	50	53	52	52	48	46	50.6	3.2	2.3	6.4

Freq : Frequency of the radio transmitter of each animal // MCP : Minimum Convex Polygon
MCP95 : Minimum Convex Polygon, based on 95% of the observations // HR95 : the home range model developed for this PhD.
Core area : area within the HR95 where the use-intensity is higher than the average use-intensity within HR95
%area: expresses the home range model mentioned directly above as a percentage of the MCP
%obs: the number of fixes within the home range definition used (directly above) expressed as percentage of the total number of fixes.
Cv. :coefficient of variation.

The results show a large variation in the size of the MCP, MCP 95, HR95 and the core area (CA) with Coefficients of variation of 66.4%, 71.6%, 76.6%, and 80.2% respectively. In contrast to the large variation in the size of the home ranges, there was only a very limited variation in the size of the MCP95, HR95 and core areas when expressed as a percentage of the MCP. The same limited variation was found for the percentage of the animal fixes situated in the HR95 and the core area.

VI.E.4 Habitat selection

Because of the limited number of animals compared to the number of habitat types we were forced to reclassify the 21 habitat types in five main types to study the habitat selection on second and third order. These five classes were: forest, shrub vegetation, heathland (and grassland), open areas and areas of human activity. The H_0 (use is proportional to availability) was rejected both on second order ($p < 0.001$) and on third order of selection ($p < 0.05$). However, few significant differences in preferences for the different habitat types could be found on the second or third level. Furthermore, some habitat types could be reclassified into two different main types. This resulted in different p-values for the rejection of the H_0, and completely different preference rankings. Because of these results and the fact that reclassifying the habitat types is a completely subjective decision, we decided not to investigate the use of the different habitat classes, based on vegetation type, any further. However, based on the rejection of the H_0, one can conclude that the roe deer fawns do not use the different habitat types in proportion to their availability, and consequently, that they prefer certain habitat types over others.

VI.E.4.a Second order selection for habitat edges

Table VI.3 shows for each of the possible use/availability combinations the probabilities of the Null hypothesis in the compositional data analysis at the second level. The Null hypothesis is that the roe deer fawns use the different distance zones (from the nearest edge) randomly, or proportional to their prevalence. Rejecting H_0 means that there is a selection by some individuals for certain buffer zones, without automatically indicating that the preferences are the same for all individuals. The results show that using HR95 as the definition of used area gives the most consistent answers, both for the rejection level of H_0 and for the significance of the relative preferences and avoidances of the different distance zones. The analysis using HR95 is the only one that shows, for all the three possible definitions of available habitat, a highly significant avoidance for zones more than 45 meters from the edges. With the exception of the analysis using the composition of the CA as 'use', all analyses return, whenever the Null hypothesis is rejected, the same ranking for the preference of the different distance zones. The areas closest to the edge are preferred over the more distant areas.

TableVI.3: Preference for habitat edges (buffer zones) at the second level of selection (home range versus study area), using different methods to define 'use' and 'availability'.

Availability	Study area						MCPtot						Sommcp					
Use	p = 0.11	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking	p = 0.18	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking	p = 0.0597	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
MCP	d > 45m						d > 45m						d > 45m					0
	45m > d >30m						45m > d >30m						45m > d >30m	++				1
	30m > d > 15m						30m > d > 15m						30m > d > 15m	+	+			2
	15m > d						15m > d						15m > d	+	+	+		3

Availability	Study area						MCPtot						Sommcp					
Use	p = 0.058	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking	p = 0.075	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking	p = 0.038	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
MCP95	d > 45m					0	d > 45m					0	d > 45m					0
	45m > d >30m	+++				1	45m > d >30m	+++				1	45m > d >30m	+++				1
	30m > d > 15m	+++	+			2	30m > d > 15m	+++	+			2	30m > d > 15m	++	+			2
	15m > d	+++	+	+		3	15m > d	++	+	+		3	15m > d	++	+	+		3

P-value: the probability of accepting the H_0 hypothesis (= no preference).

Whenever $p < 0.1$ the preference matrix is generated. A positive sign indicates that the buffer mentioned in the row is preferred over the one in the column. (+ : preference though not significant / ++ : significant at 0.1 / +++ significant at 0.05). Negative signs have exactly the opposite meaning

MCPtot : MCP based on the pooled locations of all animals together

Sommcp : sum of the MCP's of the individual animals.

Table VI.3: Continued

Availability
Use

HR95

Study area

p = 0.039	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
d > 45m					0
45m > d >30m	+++				1
30m > d > 15m	+++	++			2
15m > d	+++	+	+		3

MCPtot

p = 0.054	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
d > 45m					0
45m > d >30m	+++				1
30m > d > 15m	+++	++			2
15m > d	+++	++	+		3

Sommcp

p = 0.043	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
d > 45m					0
45m > d >30m	+++				1
30m > d > 15m	+++	+			2
15m > d	+++	+	+		3

CA

Study area

p = 0,057	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
d > 45m					0
45m > d >30m	++				1
30m > d > 15m	+	+			3
15m > d	+	+	-		2

MCPtot

p = 0,068	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
d > 45m					0
45m > d >30m	++				1
30m > d > 15m	+	+			2
15m > d	+	+	+		3

Sommcp

p = 0,031	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
d > 45m					0
45m > d >30m	++				1
30m > d > 15m	+	+			3
15m > d	+	+	-		2

VI.E.4.b Third order selection for habitat edges

Tables VI.4: Preferences for the buffer zones at the third level of selection (internal use versus home range).

Availability
Use

HR95

CA

p = 0.087	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
d > 45m					0
45m > d >30m	+				2
30m > d > 15m	+	+			3
15m > d	+	-	-		1

MCP

CA

P = 0.0969	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
d > 45m					0
45m > d >30m	+++				1
30m > d > 15m	+	+			3
15m > d	+	+	-		2

MCP

Obs. MCP

p = 0.093	d > 45m	45m > d >30m	30m > d > 15m	15m > d	Ranking
d > 45m					0
45m > d >30m	+				2
30m > d > 15m	+	-			1
15m > d	+	+	+++		3

P-value: the probability of accepting the H_0 hypothesis (= no preference). Whenever $p < 0.1$ the preference matrix is generated. A positive sign indicates that the buffer mentioned in the row is preferred over the one in the column. (+ : preference though not significant / ++ : significant at 0.1 / +++ significant at 0.05). Negative signs have exactly the opposite meaning.

On the third order of selection the results are less convincing. Comparing the fixes in the HR95 with the availability in the MCP, MCP95 or in HR95 never resulted in rejecting the H_0 hypothesis of random use. Using the CA as input for the used area within the home range resulted in rejecting the H_0 hypothesis of random use when the MCP or HR95 were used to define the availability. Because some of the observations of the MCP are located outside the MCP95 and the HR95, we could only compare the locations in the MCP with the availability in the MCP. H_0 is rejected for this comparison. When comparing the three cases were the H_0 hypothesis is rejected, the preference rankings are contradicting each other (Table VI.4). The only clear result in each analysis is the negative selection for zones more than 45 meters away from the nearest transition zone.

VI.E.5 Habitat characteristics

When comparing the home ranges of the eight roe deer fawns with the simulated ones, the Mann-Whitney test indicated a significant difference ($p < 0.05$) for the number of patches (NP), the patch density (PD) and the edge density (ED). There was no difference between the total area (TA) or the total edge length (TE) (Table VI.5).

Similar conclusions resulted from the randomisation tests (Figure VI.3 and Table VI.5). The probability that mean of the set of the eight true roe deer home ranges belongs to the same population as the means of the 1000 random sets is below 0.001 for ED, NP and PD. The p-values for TA and TE are respectively 0.44 and 0.16. The animal home ranges are therefore characterised by a significant higher ED, NP and PD.