

Density Estimation and Time Trend
Analysis of Large Herbivores in Nagarhole,
India

Aditya Gangadharan
Mres in Environmental Biology
University of St. Andrews
2005

Supervisor: Dr. Len Thomas

ACKNOWLEDGEMENTS

I am grateful to my supervisor Dr. Len Thomas for his generous and cheerful advice throughout the course of this project. I thank Mr. N Samba Kumar and Dr. K Ullas Karanth for making this project possible and for their valuable advice and suggestions. The code used for the bootstrap routine was based on that written by Tiago Marques.

Contents

	Page
Abstract	4
1. Introduction	1
1.1 Background	1
1.2 Distance Sampling Theory	3
1.3 Trend Modelling	5
2. Methods	8
2.1 Exploratory Analysis and Forward Truncation	8
2.2 Density and Confidence Interval Estimation	11
2.3 Trend Modelling	14
3. Results	17
3.1 Forward Truncation Distances	17
3.2 Model Selection and Density Estimates	29
3.3 Trend Analysis	35
4. Discussion	41
4.1 Density Estimation	41
4.2 Trends	45
4.3 Improvements, Recommendations and Extensions	46
References	51
Appendix 1: Forward Truncation for Sambar and Elephant	54
Appendix 2: Quantile-quantile Plots by species	57
Appendix 3: Density Estimates by Habitat Stratum	63
Appendix 4: Trends by habitat stratum for some species	70
Appendix 5: R code for Bootstrap	73

ABSTRACT

Density estimates for six large herbivore species were obtained through analysis of line transect data from Nagarhole National Park, south-western India, collected between 1989 and 2000. These species were Chital (*Axis axis*), Sambar (*Cervus unicolor*), Gaur (*Bos gaurus*), Wild Pig (*Sus scrofa*), Muntjac (*Muntiacus muntjak*) and Asian Elephant (*Elephas maximus*). Multiple Covariate Distance Sampling (MCDS) models were used to derive these density estimates. The distance histograms showed a relatively large spike at zero, which can lead to problems when fitting MCDS models. The effects of this spike were investigated and remedied by forward truncation. Density estimates from unmodified dataset were 10-15% higher than estimates from the forward truncated data, with this going up to 37% for Muntjac. These could possibly be over estimates. Empirical trend models were then fit to the density estimates. Overall trends were stable, though there were intra-habitat differences in trends for some species. The trends were similar both in cases where forward truncation was done as well as in those where they were not.

1 INTRODUCTION

1.1 Background

Monitoring long-term trends in large herbivore populations is an essential part of wildlife management; however, this has not been a priority in tropical Asia (Karanth & Stith 1999). In this context, therefore, the almost annual surveys of large herbivores in Nagarhole by the Wildlife Conservation Society (India) gains importance. This study analyses the results of these surveys and aims to contribute towards filling in this gap.

Nagarhole National Park is located in the state of Karnataka in the south-west of India. It covers an area of 644 square kilometres, and is an important refuge for several wildlife species, with an especially large assemblage of ungulates (Karanth et al 1999; Karanth & Sunquist 1995). These large herbivores include Chital (*Axis axis*), Sambar (*Cervus unicolor*), Gaur (*Bos gaurus*), Wild Pig (*Sus scrofa*), Muntjac (*Muntiacus muntjak*) and Asian Elephant (*Elephas maximus*). Apart from being important in themselves, these species are important as a source of food for carnivores. Karanth & Stith (1999) determined that prey availability was a critical requirement for viable tiger populations, and could possibly explain the lack of tigers even in areas with relatively low tiger poaching. This concurs with the findings of Ramakrishnan et al (1998). Karanth et al (2004) were also able to predict tiger density as a function of prey density. These species are therefore critical to the functioning of the ecosystem as a whole (Owen-Smith 1987).

Nagarhole is covered mostly by tropical moist deciduous forest which gives way to dry deciduous forests in areas with less precipitation (Madhusudan & Karanth

2000, p.342). This forms a gradient from the west to the east of the park. Many of the moist deciduous forests were cleared to make way for teak plantations between 1870 and 1980; this has now given rise to a distinct habitat characterised by Karanth & Sunquist (1992) as 'Teak Dominant Forests'. This habitat consists of a "mosaic of natural moist forest and plantations of teak" (Karanth & Sunquist 1992). For the purposes of this project, the above three will be referred to as 'habitat strata'.

Within these habitat strata, the above authors also identified microhabitat types: artificially maintained grass clearings and swamp fallows called hadlus. In addition, patches of secondary forests occur where past teak plantations failed (Madhusudan & Karanth 2000); where they have not, the patches are referred to as 'teak plantations'. Thus at the microhabitat level, six habitat types were identified: Dry Deciduous Forest, Moist Deciduous Forest, Teak Plantation, Secondary Forest, Clearing and Hadlu. These are referred to as 'habitat types' in this study. The dry deciduous habitat stratum consists mostly of dry deciduous habitat types; this is true to a lesser extent for the moist deciduous stratum as well. The highest diversity of habitat types occurs in the teak dominant stratum.

This study analysed data collected in the 104 square kilometre Nalkeri Reserved Forest in the west of Nagarhole. 8 line transects were established in the late 1980's with a total length of 18 kilometres. These transects were laid purposively and sampled the habitat strata in proportion to the areas covered by them: thus 30% of the transect length was in dry deciduous forest, 30% in moist deciduous forest and 40% in teak dominant forest. Figure 1.1 shows the map of the Nagarhole with transect locations (A digitised map of Nalkeri was not readily available, but the transects roughly represent its outline).

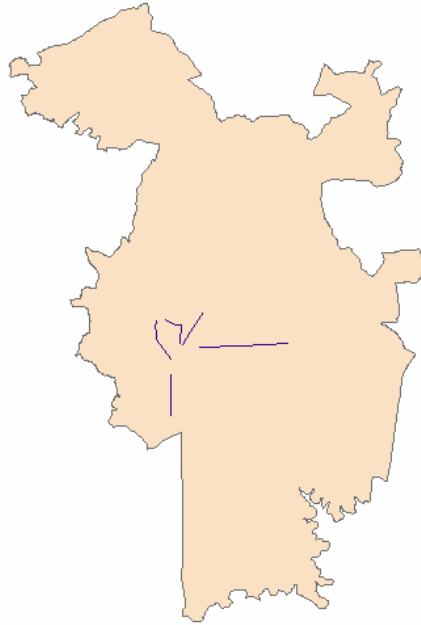


Figure 1.1: Map of Nagarhole with transects marked

These transects were surveyed on a minimum of 26 occasions in each year, with these temporal replicates carried out during a period of around 15 days each year. This study analysed 10 years of data collected between 1989 and 2000 (no surveys were carried out in 1992 and 1993). Data was analysed using the statistical software R version 2.1.1 (R Development Core Team 2005) and Distance 5 Beta 4 (Thomas et al 2005).

1.2 Distance Sampling Theory

Line transect are a form of distance sampling where the observer traverses a line, searching for animals in a 180° arc in front and to either side upto a distance w (animals observed behind the observer may also be recorded). The observer generally records the sighting distance and angle of the observed animal cluster, from which perpendicular distances of these clusters from the transect are calculated by trigonometry. A detection function curve $g(x)$ is fitted to the histogram of these

perpendicular distances. All animals on the line are assumed to have been seen ($g(0)=1$). This graph is then rescaled to give a probability density function with the same shape as the detection function but so that it integrates to 1. The probability of detection is obtained by dividing the area under the curve by w . Thus the estimated probability of detection \hat{P}_a and estimated density \hat{D} are given by (Buckland et al 2001, pp.51-53):

$$\hat{P}_a = \frac{1}{w} \int_0^w g(x) dx \quad (1)$$

$$\hat{D} = \frac{\hat{E}(s)n}{2wL\hat{P}_a} \quad (2)$$

where w is the distance upto which observations are made (or the perpendicular distance data truncated), $g(x)$ is the detection function, $\hat{E}(s)$ is the estimated cluster size, n is the number of clusters observed and L is the total survey effort.

Conventional Distance Sampling (CDS) assumes that perpendicular distance from the line is the only factor affecting probability of detection. This is ensured by the use of models that are ‘Pooling Robust’ (Buckland et al 2001, p. 41). Pooling robustness ensures that “heterogeneity in the detection probabilities due to covariates other than distance are ignored” (Marques & Buckland, 2004, p.31). Thus global estimates derived from a global detection function should be quite similar to those derived by fitting separate detection functions for each stratum. This can, however, cause bias if a global detection function is used to estimate stratum specific densities. To get around this, detection probability may be estimated separately for each stratum, or stratum specific covariates other than perpendicular distance may be included in modelling the detection function. The latter option was chosen for this project since they were readily available, and it would also be useful for the purposes

of future surveys to identify the different factors that affected detection probability. It would be especially interesting to fit the year as a covariate since this would give better estimates of year specific density, which would be useful in the trend modelling.

The inclusion of covariates other than perpendicular distance is only possible in the half normal and the hazard rate models. The covariates are assumed to affect the detection probability through the scale parameter. Thus the detection probability $P_a(\underline{z}_i)$ of cluster i given covariates \underline{z}_i is (Marques and Buckland 2004, p.38):

$$\hat{P}_a(\underline{z}_i) = \frac{1}{w} \int_0^w g(x, \underline{z}_i) dx \quad (3)$$

where $g(x, \underline{z}_i)$ is the detection function involving both perpendicular distances as well as covariates \underline{z}_i where i labels the different covariates. From this, animal densities are obtained through a Horvitz - Thomson like estimator:

$$\hat{D} = \frac{1}{2wL} \sum_{i=1}^n \frac{s_i}{\hat{P}_a(\underline{z}_i)} \quad (4)$$

(Marques and Buckland 2004, p.41).

More details on distance sampling may be obtained from Buckland et al (2001) and Buckland et al (2004).

1.3 Trend Modelling

The purpose of this exercise was to estimate preliminary population trends over time, which had not been done so far in Nagarhole. Several methods exist for estimating trends over time (Thomas 1996); for this analysis, Generalised Additive Models (GAMs) were chosen. GAMs express the response variable as a smooth

function of the explanatory variables; they are non parametric and do not assume linear relationships between the response and the explanatory variables (Hastie & Tibshirani 1990, p.9). This makes them useful for general description and exploratory analysis of trends (Hastie & Tibshirani 1990, p.9). This tied in well with the purpose of this analysis.

In a GAM (as in a Generalised Linear Model) the response is related to the explanatory variables through a ‘link function’. The mean of the response variable is a smooth function of the additive predictor. Assuming different trends in different habitats, and a log link, estimated density \hat{D}_t in year t for habitat stratum i is modelled by:

$$\hat{D}_{it} = \exp(\beta_i + s_i(t)) + \varepsilon_{it} \quad (5)$$

where i labels each different habitat stratum, t labels the years, β_i is the intercept parameter for habitat stratum i , s_i is a smoothing function of time t and habitat i and ε_{it} is the combined process and sampling error which is independent between years and Poisson distributed in this case. The response is assumed to be from a member of the exponential family (Hastie & Tibshirani 1990) which include the normal, Poisson, binomial and other commonly used distributions. Thus a GAM involves estimation of the intercept parameter as well as the smoothing function.

The amount of smoothness to allow is a model selection question; a linear model gives maximum smoothness, whereas minimum smoothness is obtained by exactly fitting each data point (Thomas et al 2004). This is expressed in terms of the degrees of freedom for a model- the higher the degrees of freedom the more wiggly the curve. Choosing the number of degrees of freedom is not an exact science, and depends on the purposes of the analysis; Fewster et al (2000) recommend starting

with a small number of degrees of freedom and increasing it until further increases “serve only to roughen the output”. An alternative is to use automated selection techniques such as penalised regression splines or smoothing splines (Hastie & Tibshirani 1990), which is what was done for this exercise.

More details on generalised additive models may be found in Hastie & Tibshirani (1990).

2 METHODS

2.1 Exploratory Analysis & Forward Truncation

Initial analysis of the perpendicular distance data showed a relatively large number of observations at small perpendicular distances. The extent of this problem varied between species. Such observations give rise to “spiked” distance histograms, which violate the shape criterion of distance sampling (detection probability should not immediately drop off at small perpendicular distances away from the line, Buckland et al 2001 p.42). When modelling the detection function the analyst can either attempt to fit the “spike” using the hazard rate key function, or use the half-normal which will not fit it. These lead to the following issues:

- (a) The hazard rate model is sensitive to rounding of angles to zero, which may lead to large overestimates should it attempt to fit a spurious spike (Buckland et al 2001, p.43). When fitting covariates other than distance, this choice also leads to convergence problems and unstable parameter estimates.
- (b) If the spike is thought to be an artefact of the survey process, an alternative is to fit a half normal model which will fit through the spike. However, this can bias the model selection process; AIC may favour covariates that fortuitously explain the spike (giving a higher intercept) when these covariates do not in fact have much effect on the probability of detection.
- (c) It is also possible that the spike represents the true detection function, in which case models that do not fit it will underestimate abundance.

The first part of the analysis focussed on trying to understand the reasons for the spike, and possible solutions to it. The following were hypothesised as possible reasons:

- (a) Rounding of angles- Observers tend to round sighting angles to zero, resulting in a relatively large number of observations on the line.
- (b) “Guarding the Centreline” (Buckland et al. 2001, p. 30)- Observers tend to concentrate on detecting all objects on or very near the line, and pay too little attention to those slightly further away. The number of observations made and hence the detection function fall steeply with distance from the line. The histogram therefore reflects the true shape of the detection function, but is difficult to model as explained above.
- (c) High visibility on straight lines- Most of the lines are straight, and visibility along the line is good for long distances. Observers therefore see and record animals on the line far ahead of them (Thomas 2001). Many of these far sightings recorded as on the line, however, may be of animals crossing the line. This violates the assumption of distance sampling that animal movement is not important (or alternatively that the survey occurs at a “snapshot” in time, Buckland et al 2001, p. 31). Habitat heterogeneity plays an important role in this. In the more thickly forested habitats, where it is difficult to see any animal that is off the cut transect, a steep detection function with a spike at zero is likely even in normal circumstances (Buckland et al 2001, p.294). Straight transects may add to this, giving a large spike at zero.
- (d) Transects as trails- Animals may be using the transects as walking trails. Therefore animal density may not be uniform with respect to distance from the line.

It was determined that “Guarding the Centreline” was not a likely cause of the problem, since the observers are trained to avoid this. Animal use of the transects as trails was restricted to elephants possibly using one transect. Therefore the exploratory analysis focussed on investigating the effects of reasons (a) and (c) above.

Histograms of the perpendicular distance data and sighting angles were examined at various scales to investigate the effect of rounding of sighting angles. The effects of straight lines was studied by plotting the perpendicular distance of each observed cluster against its distance along the line in front of the observer (z). z was obtained by multiplying the cosine of the sighting angle by the sighting (radial) distance. It was hypothesised that if the spike was caused largely due to straight lines, a large number of observations made on the line would appear at large z distances. This would be more pronounced in the habitats with heavy undergrowth if habitat played a role in this. These analyses were carried out for the whole site as well as for each habitat stratum and habitat type. A preliminary analysis was then carried out on the full dataset for each species in Distance to check model fit and judge whether remedial measures were necessary.

Based on the results of the above analyses, forward truncation was carried out if necessary. This consisted of truncation of observations along the z distance. This is equivalent to ‘forcing’ the observer to search for animals only within the forward truncation distance specified. The aim of forward truncation was to decrease the number of observations at zero while causing as little decrease in the observations at non-zero perpendicular distances as possible. This process was essentially a trade-off between model fit and bias in density estimates. The data was forward truncated until it was deemed that the improvement in model fit caused by the decreasing the number of observations on the line would be outweighed by bias in density estimates caused

by loss of observations at non zero perpendicular distances, in particular those close to the line. If the spike at zero was associated with only with certain habitat strata or types, forward truncation was carried out only in these. Depending on the species, therefore, the data was forward truncated at the level of the whole site, habitat stratum, habitat type or not at all.

2.2 Density & Confidence Interval Estimation

Following this the data was analysed in Distance. In the species where forward truncation was carried out, both the forward truncated and untruncated datasets were analysed and density estimates compared. Analysis and model selection were carried out in accordance with the recommendations made in Thomas et al 2005, Buckland et al 2001 and Buckland et al 2004. Conventional Distance Sampling (CDS) models were fitted to the data, followed by Multiple Covariate Distance Sampling (MCDS) models with one covariate each. In accordance with the recommendations by Marques & Buckland (2004), care was taken to ensure that not more than 5% of the observations had a detection probability of less than 0.2 and none less than 0.1. This involved right truncation of between 7 and 15% of the data, depending on the species. Akaikes Information Criterion (AIC) was used to select between the different models. Covariates were added sequentially to the MCDS models, one at a time until there was no decrease in AIC. When fitting models to the forward untruncated data, it was decided not to choose models that tried to fit the spike in the species for which this was most problematic. In one case, this resulted in disregarding the model favoured by AIC.

The covariates used in the modelling were habitat stratum, habitat type, sighting time (as a factor with 2 levels- AM and PM), cluster size (continuous variable) and year (either as a continuous variable or a factor with 10 levels). Apart from these, data on the observer used on each survey was also available but not used due to the extremely large number of observers (~80) used over the ten-year period. Habitat stratum and habitat type were not used together since they were found to be highly correlated. Once the final model - which was always an MCDS model - was chosen, adjustment terms were added sequentially using the automatic selection option, setting the number of terms to a maximum of 2. Model fit was examined through goodness of fit tests given in Distance as well as Quantile-Quantile (QQ) plots and perpendicular distance histograms.

Due to technical reasons, Distance does not currently provide estimates of detection probability or density by stratum when cluster size is a covariate. Therefore these were calculated in R using the parameter estimates provided by Distance using equations (3) and (4) given in section 1.2. Density estimates were calculated for each habitat stratum, and the overall density in the study area for each year was determined by weighting the density estimates for each stratum by its area (modified from Buckland et al 2001, p.89):

$$\hat{D}_t = \frac{\sum_v A_v \hat{D}_{vt}}{A} \quad (6)$$

where v labels each stratum, t labels each year, \hat{D}_t is the estimated overall density in the study area in year t , \hat{D}_{vt} is the estimated density in stratum v in year t , A_v is the area of stratum v and A is the area of the entire study area. Density estimates from the full dataset were compared with estimates from the forward truncated dataset for each habitat stratum as well as overall.

Density estimates for different years are unlikely to be independent due to autocorrelation. This arises due to the fact that the same transects are surveyed, and also the global detection function that is fit for the data. Analytic variance estimators perform best when there are a large number of transects, which is not the case here. Therefore 95% confidence intervals were obtained through a non-parametric bootstrap procedure carried out in R. Bootstraps for line transects normally work by resampling between lines; however, this was not possible in this case due to the low spatial replication of transects (8 lines). However, these transects were walked a number of times in a year, and these temporal replicates were taken as the unit for resampling. The following explanation is modified from Borchers et al 2004 (p. 146). For each bootstrap replicate i ($i=1\dots B$) within a year j and habitat stratum k , L samples were chosen with replacement from the temporal replicate surveys conducted, where L is the number of such surveys in each year. Parameter estimates for each of these bootstrap replicate datasets were obtained from Distance and the density estimates D_{ijk} calculated in R and stored. These were then sorted, and the values at the 2.5th and 97.5th quantiles were taken as the lower and upper bounds respectively of the 95% confidence interval. This essentially combined different lines within a habitat stratum, assuming that there are no major variations in animal density within a stratum. This also assumes that the different lines within each habitat stratum are representative of the habitat stratum as a whole. 299 bootstrap resamples were carried out for all species except the Chital, for which the number of resamples had to be limited to 100 due to the extremely slow nature of the bootstrap process for this large dataset. The bootstrap for the untruncated chital data was even slower than for the truncated data, and the program crashed three-quarters of the way through the

procedure. This could not be repeated due to time constraints, and hence the forward untruncated chital dataset does not have bootstrap confidence intervals.

2.3 Trend Modelling

GAMs were fitted using the function `gam` from the `mgcv` package, version 1.3-5 (Wood 2000) in R. This package uses penalised regression splines and the Generalised Cross Validation (GCV) score to estimate the degree of smoothness of the model (Wood 2001). To avoid undersmoothing, the number of degrees of freedom for each smooth term was restricted to a maximum of 3. The GCV score was used to select between models, with a lower GCV score indicating a better fitting model. There were some problems in choosing the distribution to use for the response: the gamma family would have been most appropriate since it can take on only continuous non-negative values. However this was not possible since due to technical reasons the function `gam` does not allow the response variable to be zero when using gamma errors. This was a problem since there were several zeros in the dataset, and would be especially so for the bootstrap densities. Using the normal family would not have been appropriate since it could give negative estimates of density. Expected counts in each year with the area of the strip as an offset could have been modelled using a quasipoisson distribution, but this might not have been an accurate indicator of trends in density over time. The latter is also more biologically meaningful and useful. It was therefore decided to model the estimated density using the quasipoisson distribution with a log link, even though density is not a discrete variable. The negative binomial distribution was also tried; however, plots of residuals against the fitted values indicated that the variance was proportional to the mean rather than to some power of

the mean which supported the quasipoisson rather than the negative binomial distribution. In any case, the model results from each of these families were almost identical.

GAMs were fitted with an interaction between the smooth term for years and the habitat stratum. This was achieved by inserting a set of dummy variables for each habitat stratum, which took on a value of 1 in that particular habitat stratum and 0 otherwise. Thus through a single GAM, different smooths at different levels could be fitted for each habitat stratum. This can be expressed by:

$$\hat{D}_{it} = \exp(\beta_i + s_i(t)) + \varepsilon_{it} \quad (7)$$

where \hat{D}_{it} is the estimated density in year t and habitat i (independently poisson distributed), β_i is the value of the intercept for habitat stratum i , s_i is the smoothing function for habitat stratum i and ε_{it} is the combined process and sampling error which is independent between years and independently identically poisson distributed.

A separate GAM was fitted for the area weighted overall density estimates with the response a smooth function only of time:

$$\hat{D}_{ov,t} = \exp(\beta_0 + s(t)) + \varepsilon_t \quad (8)$$

where $\hat{D}_{ov,t}$ is the overall density in year t (independently poisson distributed), β_0 is the intercept parameter, s is the smoothing function and ε_t are the combined process and sampling errors which are independently identically poisson distributed.

When the preliminary models were fit, it was found that often the smooth terms for trends in some habitats were not significant, and a model that fitted a single smooth for the entire site rather than different smooths for each habitats gave a slightly lower GCV score. However, the purpose of this exercise was to explore trends in different habitat strata, and it was therefore decided to stick with the model that fitted different smooths for different habitats.

The density estimates for each year and habitat are not independent, since they come from the same transects each year fitted with a global detection function. This autocorrelation can worsen model fit since the model assumes that they are independent. This can also bias analytic confidence intervals. There is little one can do about the former at a relatively simple level; the latter, however can be overcome by estimating variance and confidence intervals through a non-parametric bootstrap. The process of obtaining bootstrapped 95% confidence intervals for the GAMs was integrated with the bootstrap density confidence interval estimation and carried out as recommended by Thomas et al (2004). The same GAM parameters used for the original models (habitat stratum density model and overall density model) were fitted to the density estimates obtained in each resample. The 2.5th and the 97.5th quantile values of the densities predicted by the GAM at each time point were then taken as the 95% confidence intervals of the fitted GAM at that point. Uncertainty in the number of degrees of freedom was incorporated into the bootstrap by allowing it to select the number of degrees of freedom for each resample, with a fixed upper limit of 3. The number of degrees of freedom chosen for each smooth term in a resample, along with whether it was significant or not, was also stored. The 2.5th and 97.5th quantiles of the degrees of freedom chosen thus represent the 95% confidence intervals for the number of degrees of freedom for each smooth term

3 RESULTS

3.1 Forward Truncation Distances

Forward Truncation by Species

Note: In the histograms of perpendicular distances, perpendicular distances are often not shown upto the right truncation distance to keep the figures compact.

Muntjac & Pig

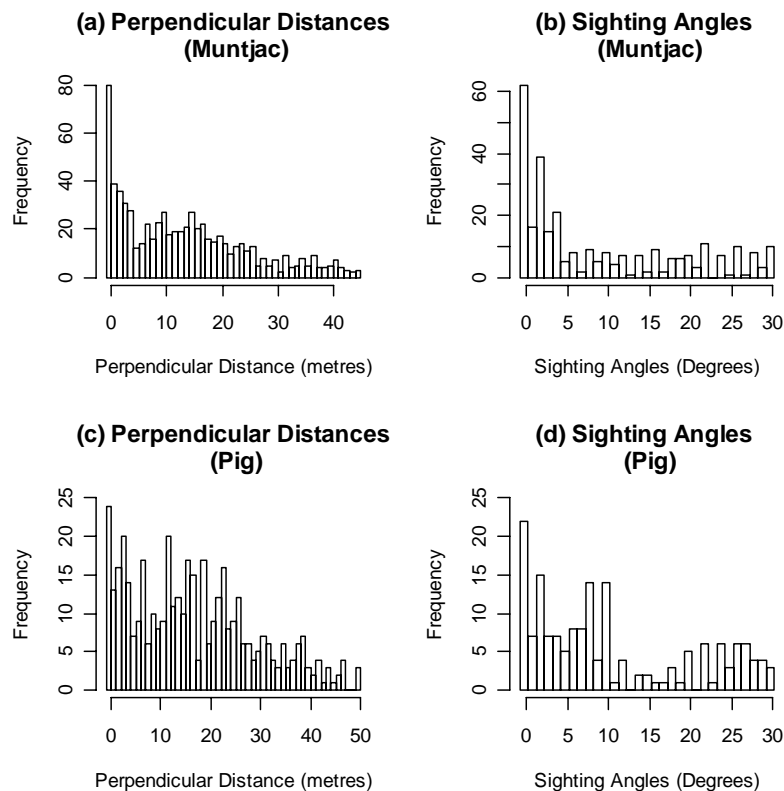


Figure 3.1a-d :distance histogram and histogram of sighting angles for muntjac and pig

The histograms of perpendicular distances (figure 3.1a and 3.1b) are fairly steep for both these species. For muntjac there is a large excess of observations on the line, whereas for the pig this is not so. The histogram of sighting angles for muntjac (figure 3.1c) indicates substantial rounding to zero, as well as rounding to the nearest

even number. This is present in a less severe form in the pig as well. This is the likely cause of the many crests and troughs in the distance histograms, such as the one around 5 metres for muntjac and at around 10 metres for the pig. This rounding of sighting angles to the nearest even number occurs to some extent in all species.

Figure (3.2a-d) shows the plots of perpendicular distance against the distance ahead of the observer for the whole site as well as for each habitat stratum for muntjac. It appears that observations on the line are made further ahead of the observer than those observations at larger perpendicular distances. This problem seems especially severe in the moist deciduous forest; 16% of the observations in this habitat are recorded as zero, compared to around 7% in the other two habitat strata. Several of the observations at zero occur beyond about 50 metres. However, there are also a large number that occur at smaller distances ahead of the observer. These are likely to be due to rounding of sighting angles, and, in the moist deciduous forests, to natural tendency for a relatively large proportion of observations to be made on the line.

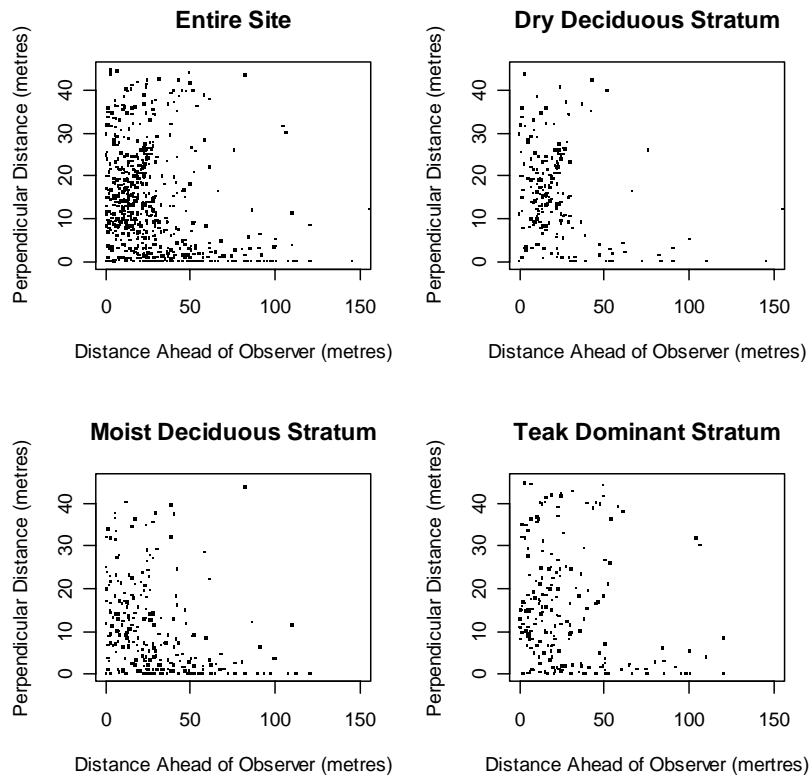


Figure 3.2 a-d : Plots of perpendicular distance against distance ahead of the observer for various habitat strata and entire site for muntjac

Similar plots by habitat type are shown in figure 3.3a-f.

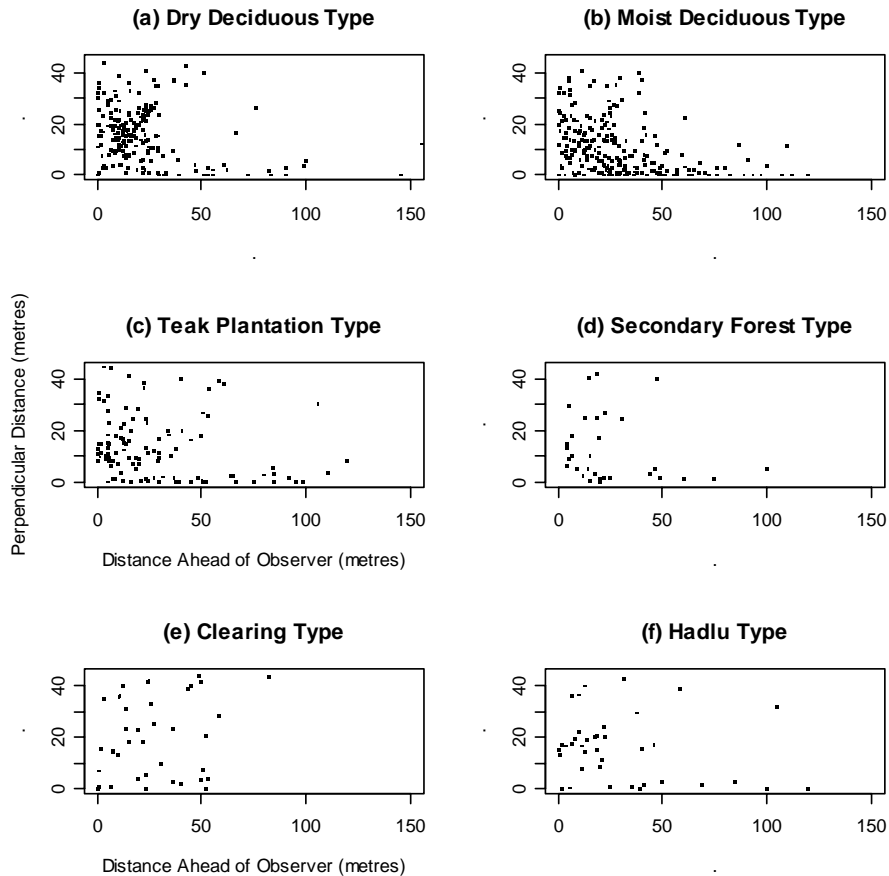


Figure 3.3 a-f : Plots of perpendicular distance against distance ahead of the observer for various habitat types for muntjac

There is clearly a contrast between observations made in clearings and in other habitats - there are few observations on the line beyond 50 metres ahead of the observer in clearings, whereas in the others there is a long line of observations that stretches far ahead of the observer. The hadlu is an open habitat, and one would expect it to show similar patterns to the clearings. However, there are zeros at large distances ahead of the observer even in this- possibly a consequence of rounding of sighting angles for far off sightings near the line or observers concentrating on locating objects on the line far ahead.

Similar plots by habitat type for pig (figure 3.4a-f) show that a large proportion of the zeros come from the moist deciduous habitat type. This also

contributes to a large extent to the steep histogram of perpendicular distances since many of these observations are at small non-zero perpendicular distances.

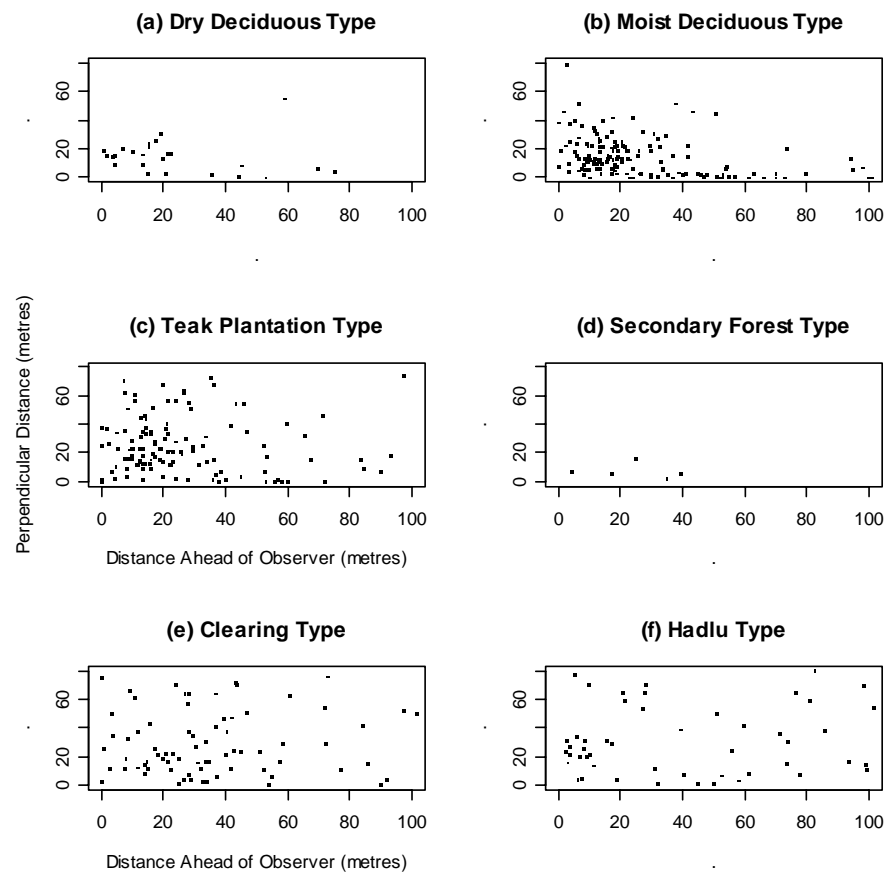


Figure 3.4 a-f : Plots of perpendicular distance against distance ahead of the observer for various habitat types for pig

Both these species have quite steep distance histograms, which is to be expected given their relatively small size. Any forward truncation must therefore not affect the steepness of the histogram close to the line, but cut down on the number of observations on the line. For muntjac it was decided to forward truncate at the level of the entire site, since this would not affect the observations in the clearings at all. It was found that truncation at distances at around 50 metres were very effective in removing the spike at zero while causing relatively small decreases in observations at other perpendicular distance close to the line. Figure 3.5a-d shows the distance histograms for muntjac at different levels of forward truncation.

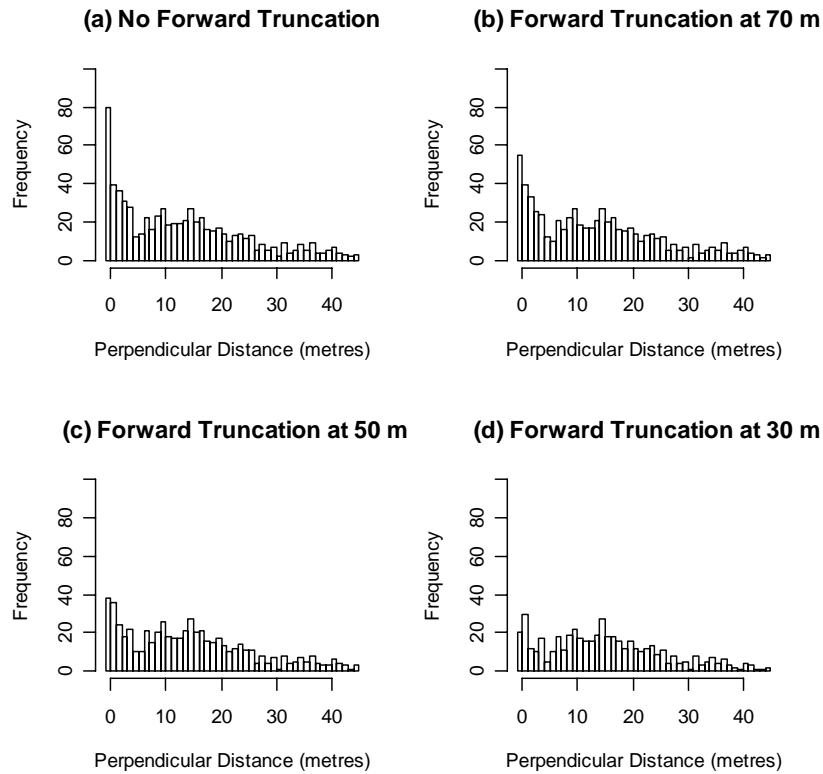


Figure 3.5 a-d: different levels of forward truncation for muntjac

For pig, it was not clear whether forward truncation was necessary; the preliminary model fitted in Distance to the untruncated data gave a non significant fit, though this was not very good. It was thought that mild forward truncation in the moist deciduous habitat type might improve the fit, and this was done at 60 metres ahead of the observer. Figure 3.6 compares the histograms in the first 20 metres for the forward untruncated and forward truncated datasets for both these species.

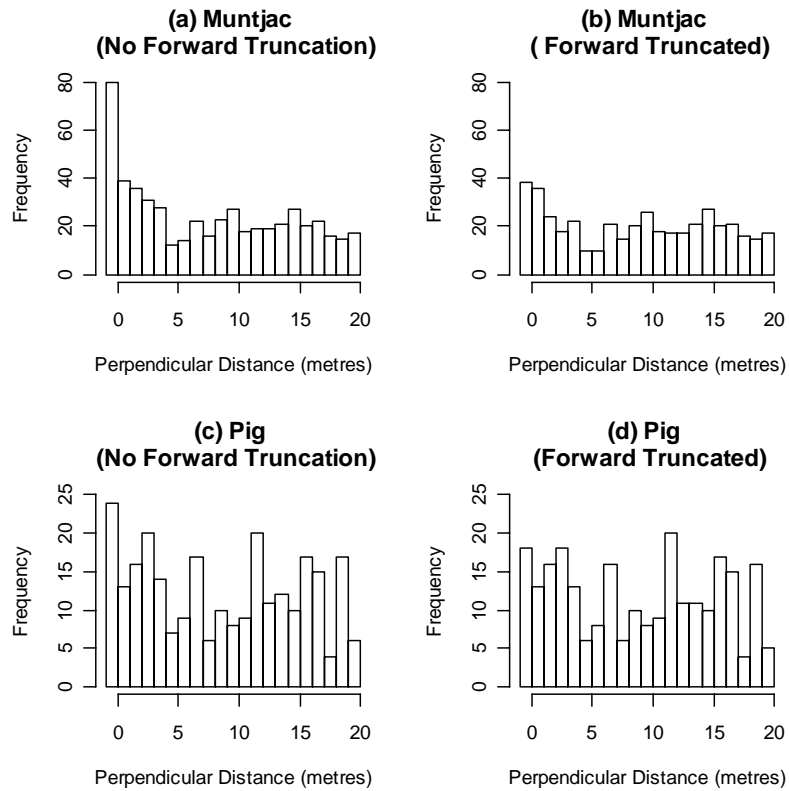


Figure 3.6 a-d: Comparison of forward truncated and untruncated datasets in pig and muntjac

Chital

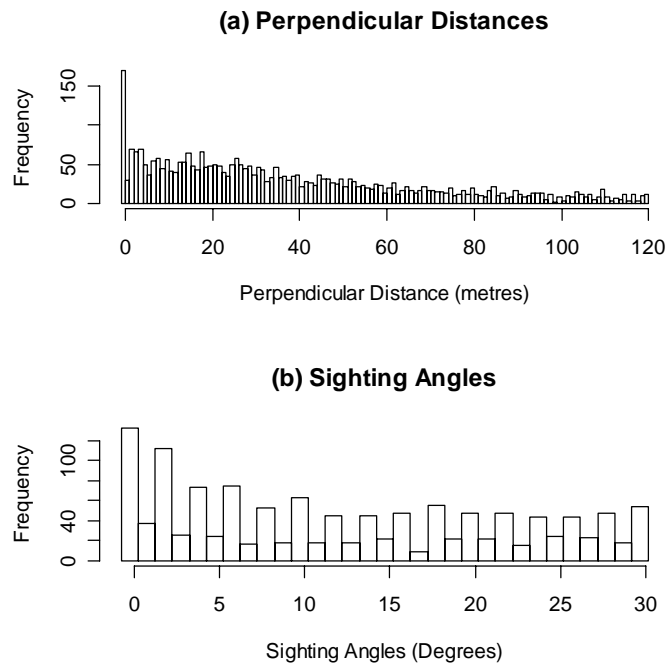


Figure 3.7 a-b: histograms of perpendicular distances and sighting angles for chital

The distance histogram for chital (figure 3.7a) shows a relatively broad shoulder if the spike at zero is ignored. There appears to be clear rounding of angles to the nearest even number (figure 3.7b). Plots of perpendicular distance against z distance by habitat type (figures 3.8a-f) show that the observations at zero are made over a large range of distances ahead of the observer in all habitats. There does not seem to be an obvious correlation between observations at zero and large distances ahead of the observer, except perhaps in the teak plantations. Only perpendicular distances in the first 40 metres are shown for clarity.

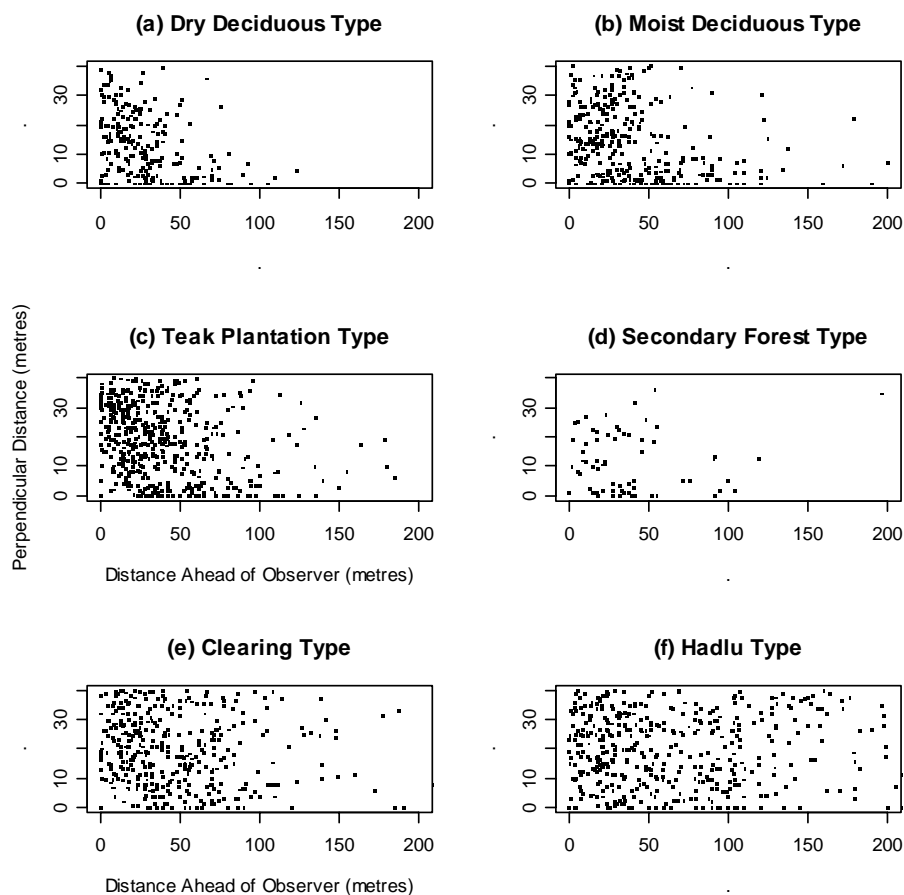


Figure 3.8 a-f : Plots of perpendicular distance against distance ahead of the observer for various habitat types for chital

The histograms of perpendicular distance by habitat type (figures 3.9a-f) indicate rounding of angles to some extent in all habitat types. There are also a relatively larger number of zeros in the forested areas compared to the open ones.

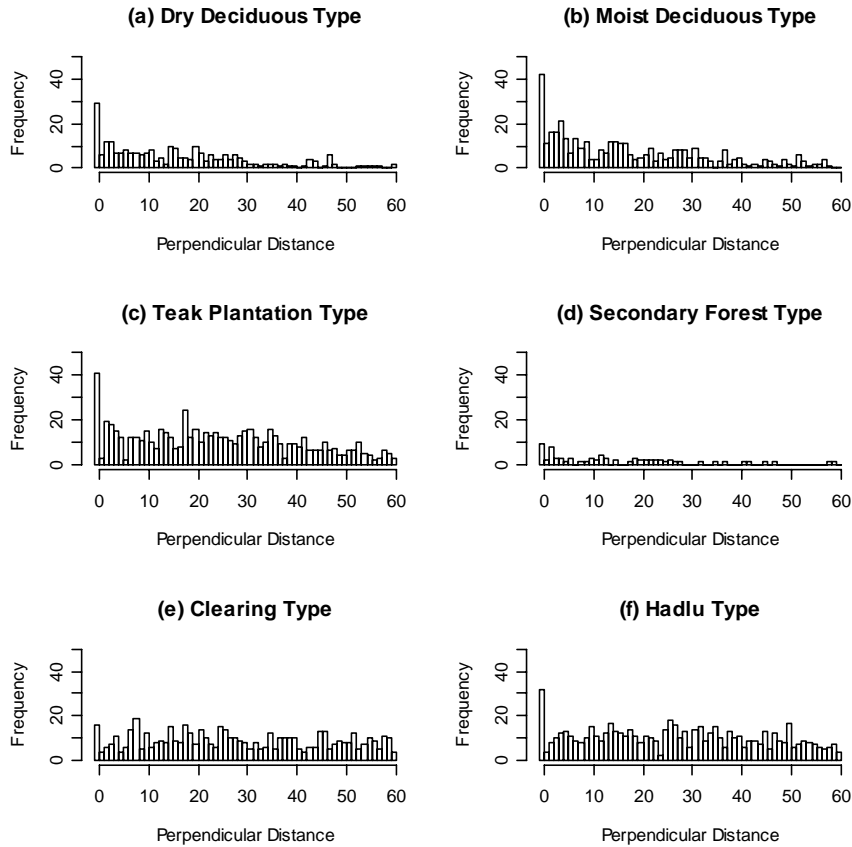


Figure 3.9 a-f : Histograms of perpendicular distance for various habitat types for chital

It was felt that rounding of angles to zero plays a large role in causing the spike in the chital; difficulty in locating these animals off the transect in forested areas leading to steep histograms was also thought to play a role. Good visibility along the line far ahead of the observer seemed to play a comparatively limited role; nevertheless forward truncation at 100 metres ahead of the observer does decrease the number of zeros without affecting other observations near the line too much. It was decided to forward truncate the entire dataset data at this distance, keeping in mind the possibility that this might not be entirely appropriate. Figure 3.10 compares the first 20 metres of the forward truncated and forward untruncated distance histograms.

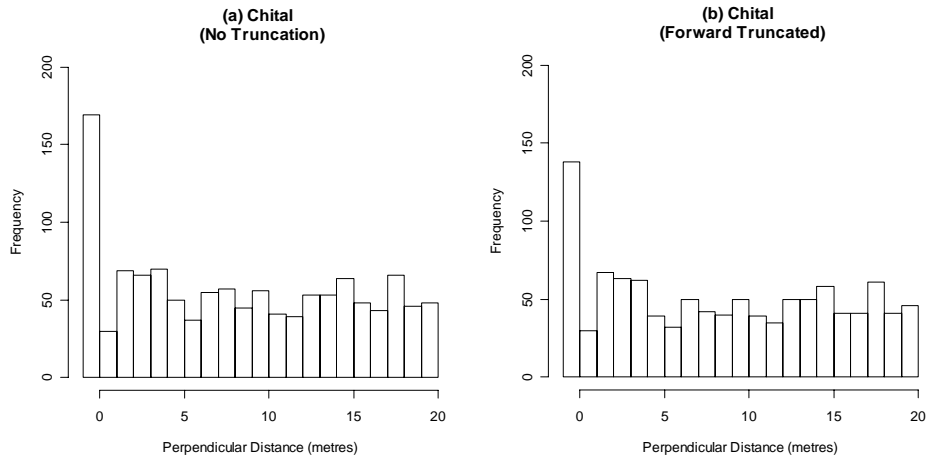


Figure 3.10 a-b: Comparison of forward truncated and untruncated datasets in chital

Sambar & Elephant

The reasoning behind forward truncation in elephant and sambar was similar to the previous examples. For both these, the distance histograms were not particularly steep except for the spike at zero. It was found that this was mostly caused in the forested habitat types (dry deciduous forest, moist deciduous forest, teak plantations and secondary forests). For the elephant, the transect that they were suspected of using as trails was not found to contribute much to the number of zeros and therefore this possibility was ignored. It was decided to forward truncate the observations in these forested habitats only. Forward truncation was carried out at 60 metres for the sambar and at 100 metres for the elephant. Figure 3.11 compares the distance histograms for these species before and after forward truncation. More details may be found in Appendix (1).

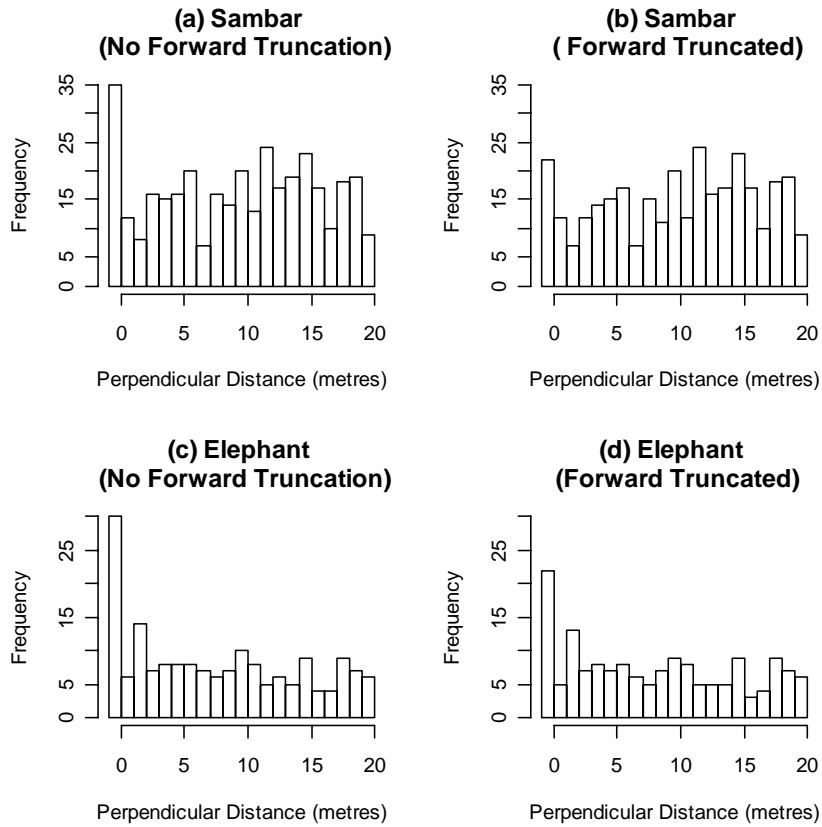


Figure 3.11 a-d: Comparison of forward truncated and untruncated datasets in sambar and elephant

Gaur

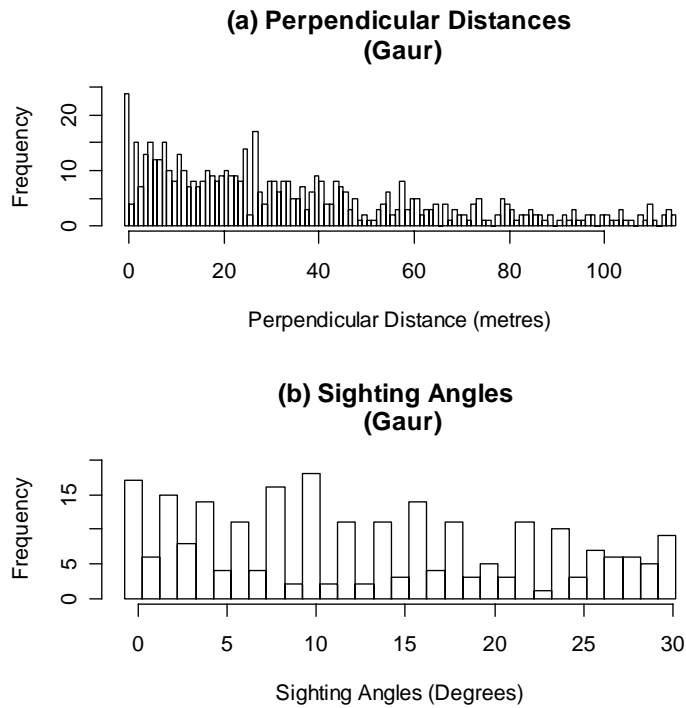


Figure 3.12 a-b: histograms of perpendicular distances and sighting angles for gaur

The spike at zero for the gaur (figure 3.12) appears to be caused to some extent by straight lines (figure 3.13), especially in the dry and moist deciduous forests. Despite this spike, however, the model fit statistics for the untruncated data in Distance were satisfactory. The many heaps and troughs caused due to rounding of angles are also likely to contribute to the worsening the fit. Forward truncation was tried at 60 metres in the forested habitat types, in which visibility along lines seems to play a role. When this was analysed in Distance, the same covariates and the same model were chosen. The density estimates from the forward truncated and forward untruncated were quite similar; therefore the forward truncated model was not pursued further and the analysis continued with the forward untruncated dataset.

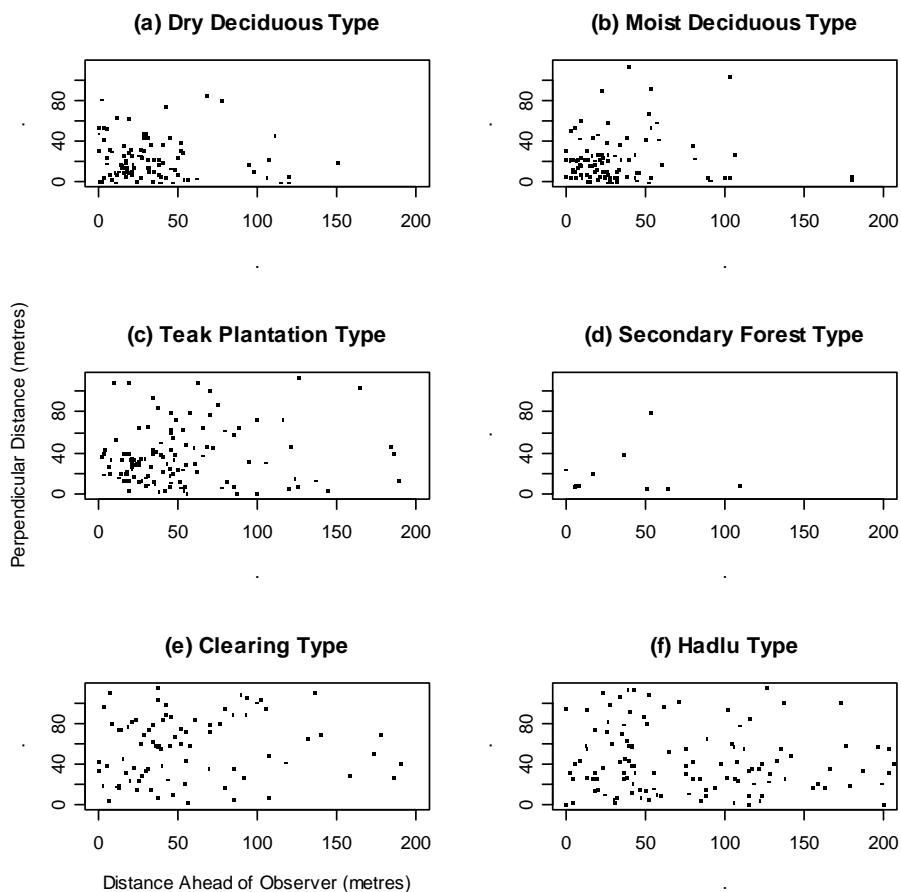


Figure 3.13 a-f : Plots of perpendicular distance against distance ahead of the observer for habitat types for gaur

Summary

The following table summarises the results of forward truncation (percentage of zeros are given as a percentage of each right truncated dataset):

Species	Forward Truncation	% of Zeros
muntjac	None	11.5
muntjac	Entire dataset at 50 m	6.5
Pig	None	5.1
Pig	Moist deciduous habitat type at 60 m	3.9
Chital	None	5.1
Chital	Entire dataset at 100 m	4.9
Sambar	None	5.7
Sambar	Forested habitat types at 60 m	3.8
Elephant	None	6.5
Elephant	Forested habitat types at 100 m	5

Table 3.1: Summary of truncation carried out by species

The biggest decrease in zeros is in the muntjac; the percentage of observations at zero in the forward truncated dataset is nearly half that of the forward untruncated dataset. The decreases for the other species are much smaller, but definite except for the chital where forward truncation seems to have made very little difference.

3.2 Model Selection & Density Estimates

Models Selected

Table 3.2 summarises the models selected for each species:

Species	Sample Size	Forward truncation	Key Function	Adjustment Terms	Covariates
---------	-------------	--------------------	--------------	------------------	------------

muntjac	696	No	Half Normal	Cosine, two	HT,Y (fac)
muntjac	587	Yes	Hazard Rate	Simple Polynomial, one	HT,Y (fac),ST
Wild Pig	473	No	Half Normal	Cosine, one	HT
Wild Pig	458	Yes	Half Normal	Cosine, two	HT
Chital	3267	No	Half Normal	Cosine, one	HT,CS,ST,Y (cont)
Chital	2816	Yes	Half Normal	Cosine, one	HT,CS,ST,Y (fac)
Sambar	615	No	Half Normal	Cosine, two	HT,CS,Y (cont)
Sambar	575	Yes	Half Normal	Cosine, two	HT,CS,ST,Y (cont)
Elephant	465	No	Half Normal	Cosine, one	HT,CS,ST,Y (cont)
Elephant	441	Yes	Half Normal	None	HT,CS,Y (cont)
Gaur	569	No	Hazard Rate	None	HT,CS

Table 3.2: Models selected for each species HT-habitat type, CS-cluster size, ST-sighting time, Y –year, fac as a factor, cont as a continuous variable

In most cases, the covariates chosen in the forward truncated and untruncated datasets are quite similar. For the elephant, sighting time was chosen as a covariate in the untruncated case whereas it was not for the truncated dataset, possibly indicating that it was chosen because it tried to fit the spike. For sambar and muntjac on the other hand, sighting time was chosen as a covariate for the forward truncated dataset

and not the untruncated one. This is at first counter-intuitive, since the forward truncated dataset has less data and one would expect fewer covariates to be fit. This could be because the forward truncated data is of better quality than the forward untruncated, and hence the covariates that really affect detection probability are chosen- perhaps this was not possible in the forward untruncated data due to the ‘noise’ caused due to the spike.

For the forward untruncated muntjac dataset, the model favoured by AIC was a conventional hazard rate model with no adjustment terms, which fit the spike to a large extent. This was ignored, and covariates were added to the model with the next lowest AIC in the manner described in section 2.2. However, even at the end of this, the conventional hazard rate model was favoured by a large difference in AIC. This was ignored, and density estimates were derived from the model that did not fit the spike even though its AIC score was around 17 higher than the CDS hazard rate one. An interesting consequence of forward truncation in muntjac is that a hazard rate rather than the half normal MCDS model was chosen for the forward truncated dataset.

Model Fit

The following table summarises the results of the Goodness of Fit tests used in determining model fit. Quantile – Quantile plots are included in Appendix (2).

Species	Forward truncation (m)	KS Test	cvM Test (uniform)	cvM Test (cosine)
muntjac	No	0	0.000 < p <= 0.001	0.000 < p <= 0.001
muntjac	Yes	0	0.010 < p <= 0.025	0.001 < p <= 0.005
Wild Pig	No	0.175	0.300 < p <= 0.400	0.150 < p <= 0.200
Wild Pig	Yes	0.2475	0.400 < p <= 0.500	0.300 < p <= 0.400
Chital	No	0	0.000 < p <= 0.001	0.000 < p <= 0.001
Chital	Yes	0	0.001 < p <= 0.005	0.000 < p <= 0.001
Sambar	No	0.0372	0.400 < p <= 0.500	0.300 < p <= 0.400
Sambar	Yes	0.3691	0.800 < p <= 0.900	0.700 < p <= 0.800
Elephant	No	0.0359	0.010 < p <= 0.025	0.005 < p <= 0.010
Elephant	Yes	0.1628	0.100 < p <= 0.150	0.050 < p <= 0.100
Gaur	No	0.2635	0.500 < p <= 0.600	0.400 < p <= 0.500

Table 3.3: Goodness of fit tests for each dataset examined. KS- Kolmogorov Smirnov test, cvM- Cramer von Mises test

The goodness of fit tests and Q-Q plots indicate mixed results for different species. In the case of the chital, forward truncation appears to have made very little difference to the model fit. It is similar for the muntjac- forward truncation has

decreased the number of observations at zero by a large amount and the Q-Q plot looks much better, but the first few metres are still not fitted very well. For pig, the model fit is now slightly better. The data for the elephant shows a significantly better fit near the line (Cv M cosine weighted test) than in the forward untruncated case, even though forward truncation removed only a few observations on the line. Also this has made a definite improvement in the largest difference between the empirical and cumulative distribution functions (KS test). For sambar the KS test is now significant, and the fit near the line is extremely good.

Density Estimates

Table 3.4 gives overall density estimates for the entire site for muntjac by year for both the forward untruncated and truncated data. Density estimates for other species as well as estimates by habitat stratum are given in Appendix (3).

Year	Forward Untruncated		Forward Truncated	
	Density (/sq km)	95% Confidence Intervals	Density (/sq km)	95% Confidence Intervals
1989	4.14	2.8 - 5.3	2.81	2.1 - 3.9
1990	3.49	2.4 - 4.9	1.52	1.1 - 2.5
1991	2.99	1.8 - 4.4	1.88	1.4 - 2.9
1994	4.73	3.1 - 6.9	2.57	1.8 - 3.7
1995	6.63	4.7 - 8.6	4.14	3.2 - 6.1
1996	6.95	5 - 9.6	4.60	3.3 - 7
1997	3.53	2.3 - 4.8	2.23	1.7 - 3.2
1998	5.13	3.5 - 6.2	3.07	2.2 - 4.2
1999	2.57	1.4 - 3.8	2.04	1.4 - 3.1
2000	3.05	1.9 - 4.4	2.06	1.4 - 3.2

Table 3.4: Density estimates confidence intervals for density in the entire site by year for muntjac

The densities estimated by the forward untruncated data are higher than those estimated by the forward truncated data in most cases. However, the point estimates are not extremely different in most species. An exception is the muntjac, where the forward truncated data gives a much lower estimate of abundance every year.

Table 3.5 ranks the species by mean abundance over the years for the entire site:

Forward Untruncated data

Species	Mean Density (/sq.km)
Chital	47.1
Gaur	6.6
muntjac	4.3
Pig	4.1
Sambar	3.1
Elephant	2.1

Forward Truncated Data

Species	Mean Density (/sq.km)
Chital	41.4
Pig	3.5
muntjac	2.7
Sambar	2.6
Elephant	1.9

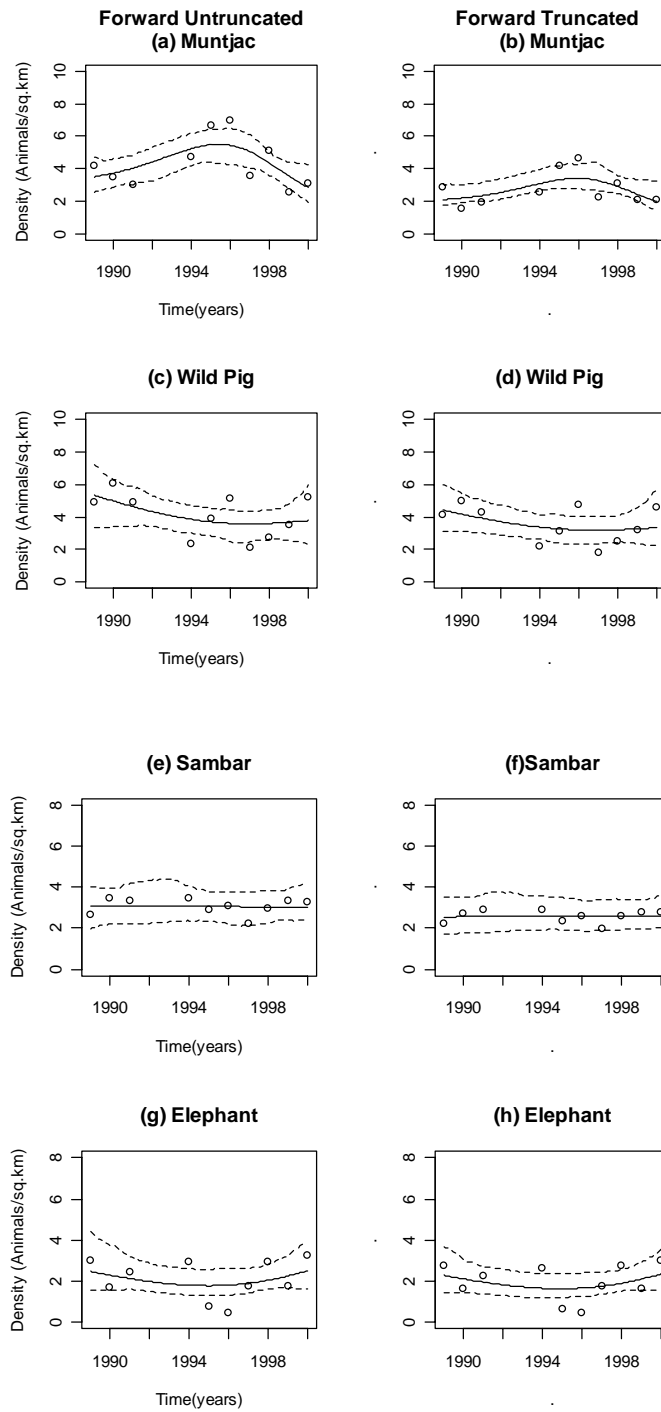
Table 3.5: Comparison of ranking of species by density with mean densities over the years analysed

The rankings in terms of density are quite similar for both forward truncated and untruncated datasets, except that the pig and muntjac exchange positions- a consequence of the large decrease in muntjac density estimates for the forward truncated dataset. The density estimates from the forward truncated data are 12% lower for chital, 15% lower for pig, 37% lower for muntjac, 16% lower for sambar and 10% lower for the elephant.

3.3 Trend Analysis

Overall Trends

The following figures show the empirical trends for each species in the site as a whole.



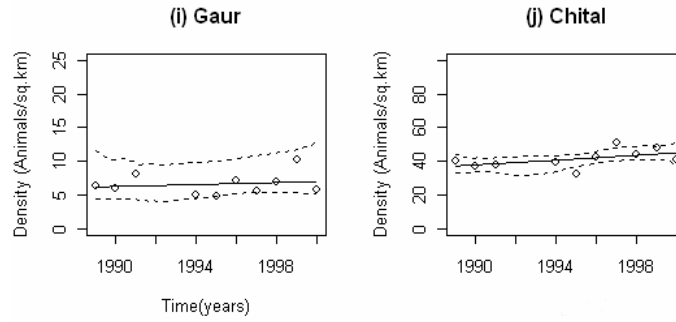
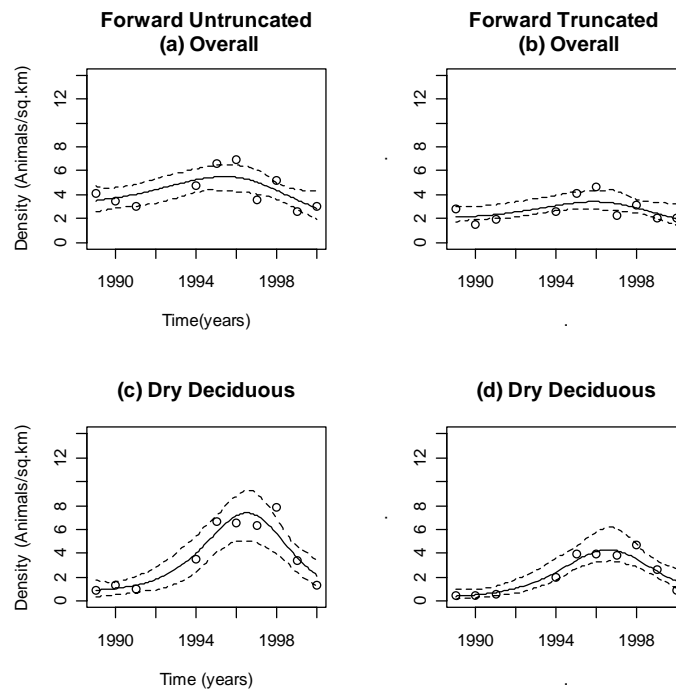


Figure 3.14 a-j: Overall trends for entire site for each species

Trends by Species

The results of trend analysis for muntjac, sambar and chital for each habitat stratum are given below. The results for the other species may be found in Appendix 4.

muntjac



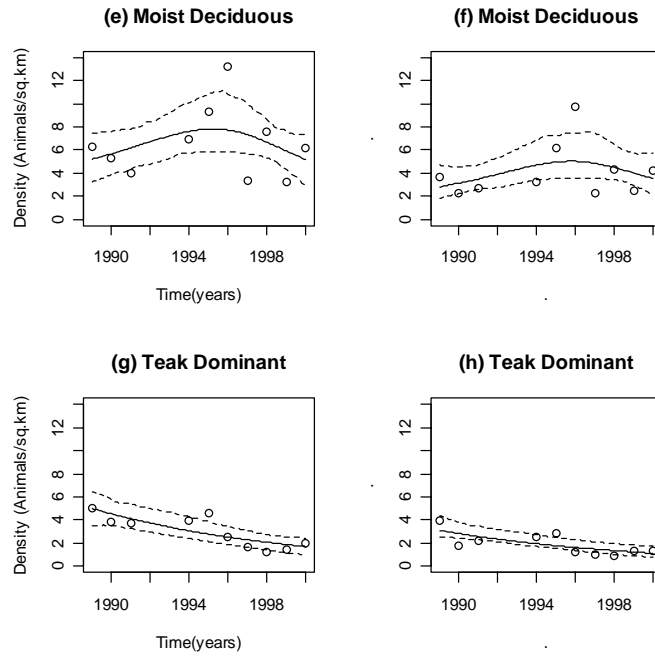
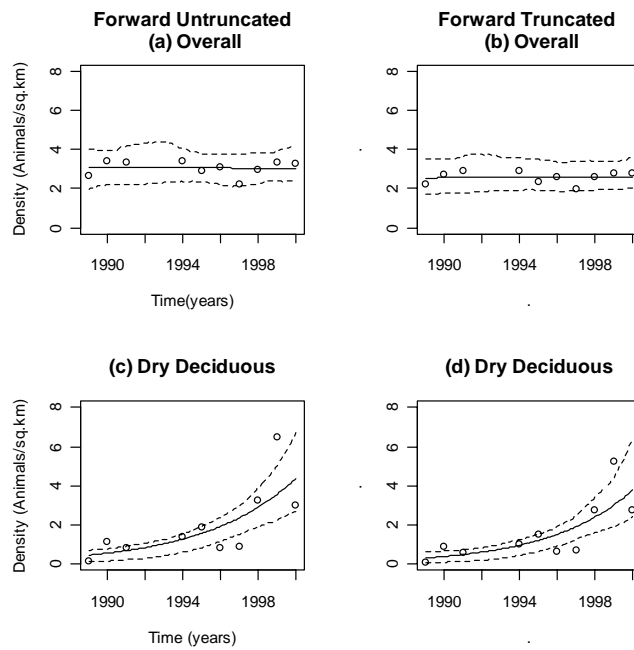


Figure 3.15 a-h: Density trends for muntjac for different habitat strata

Sambar



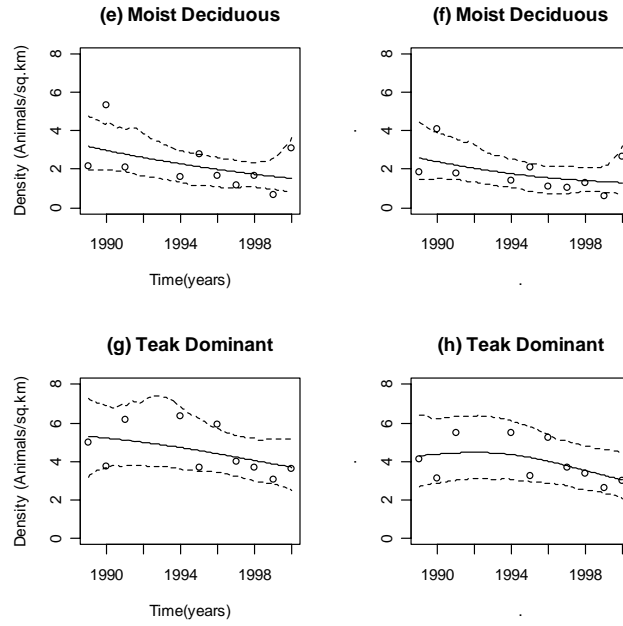


Figure 3.16 a-h: Density trends for sambar for different habitat strata

Chital (Forward Truncated)

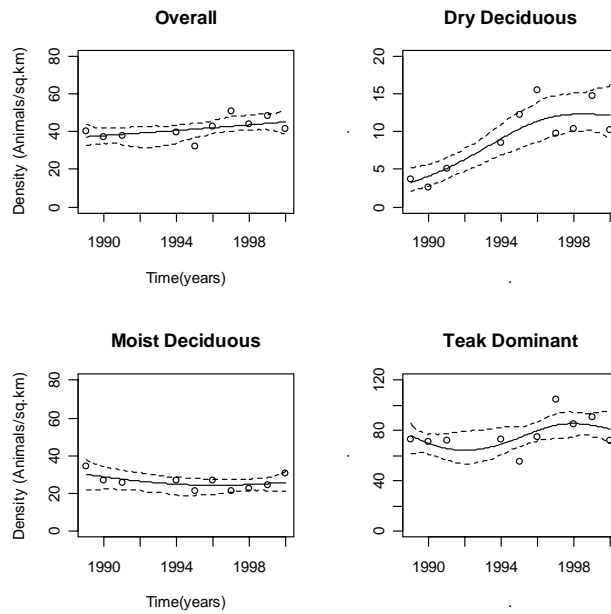


Figure 3.16 a-h: Density trends for forward truncated chital data for different habitat strata

Significance & Degrees of Freedom

The following table gives the estimated 95% confidence intervals for the degrees of freedom for each smooth, by habitat stratum as well as for the overall site. This is given for both forward truncated and untruncated data.

Forward Unruncated Forward Truncated

Species	Habitat		%Significant	Habitat	
	Stratum	CI		CI	%Significant
muntjac	Overall	1.4 - 2.8	9	1 – 2.8	4
	DD	2 - 2.9	97	1.9 - 2.8	93
	MD	1 - 2.7	13	1 – 2.7	9
	TF	1 - 1.8	44	1 – 1.4	42
Pig	Overall	1 - 2.6	8	1 – 2.6	4
	DD	1 - 1.3	17	1 – 1.3	23
	MD	1 - 2.8	36	1 – 2.9	41
	TF	1 - 1.7	16	1- 2.5	20
Chital	Overall	NA	NA	1 – 2.8	16
	DD	NA	NA	1 – 2.1	62
	MD	NA	NA	1 – 2.4	6
	TF	NA	NA	1 – 2.8	31
Sambar	Overall	1 - 2.9	7	1 – 2.9	8
	DD	1- 1.35	96	1 – 1.4	93
	MD	1 - 2.6	15	1- 2.6	10
	TF	1 - 2.5	0	1 – 2.8	3
Elephant	Overall	1 - 2.1	0	1 – 2.1	0
	DD	1 - 2.8	2	1 – 2.8	1
	MD	1 - 1.9	0	1 – 2.0	0
	TF	1 - 2.3	9	1 – 2.4	7
Gaur	Overall	1 - 2.7	3	-	-
	DD	1 - 2.7	7	-	-
	MD	1 - 1.7	38	-	-
	TF	1 - 1.6	10	-	-

Table 3.6: 95% confidence intervals for degrees of freedom chose for each bootstrap. DD- Dry deciduous, MD-

Moist Deciduous, TF- Teak Dominant Forest, NA- Not Available, - Not Applicable

For the most part, there are no definite trends in the densities over time, with most of the smooths not significant at the 5% level. In most cases, a relatively low number of degrees of freedom are fit. The exceptions to this are the muntjac, chital and the sambar densities in dry deciduous forests, where the majority of the bootstrap resamples fit significant smooth terms. A substantial number of the resamples for the teak forests for the muntjac are also significant, with low degrees of freedom, indicating linear trends if any at all. For the elephant, many of the estimated density estimates are outside the 95% confidence intervals of the GAM. This could be a result of oversmoothing. Also, there is an unusual 'bump' in the upper confidence limit in the dry deciduous forest in elephants. This is probably since density shows an upward trend until 1991, but no data is available for 1992 and 1993; when predicting the densities for these years, sometimes the GAM tries to take into account this upward trend and fit these with a smooth curve.

4. DISCUSSION

4.1 Density Estimation

Forward Truncation- Choices and Consequences

In general, the spike was found mostly in the forested areas, especially in the thickly forested moist deciduous stratum, which may be explained by straight cut transects. Hadlus show a spike in some species which is surprising; this could possibly be due to rounding of sighting angles at large distances ahead of the observer. Preferential recording of even angles by observers may be due to only even degrees being marked on compasses. On a large scale this causes heaps and troughs in the distance histogram, which makes model fit worse.

The muntjac and pig both have steep distance histograms, which is to be expected due to their relatively small size. However, the muntjac shows an extremely large spike at zero, while this is not so for the pig. This could occur if pigs tend to avoid the transects compared to other species and so there are not as many on the line as there 'should' be. The issue of the straight lines is still relevant for the pig, but since there are fewer of them on the line there may be less of a spike. For both these species, as well as for sambar and elephant, it appears that good visibility along straight lines in the forested habitats plays a large role in causing the spike; rounding of angles may be secondary.

The chital data was forward truncated at the level of the entire site, which led to the loss of around 450 observations. However, this did not make a large difference to the percentage of zeros in each dataset- 5.1% of the dataset for the forward untruncated case, and 4.9% for the forward truncated one. The former gives an

estimated overall density of around 47 /sq.km while the latter gives a lower estimate at around 41 /sq.km. It appears that forward truncation at the level of the whole site may have caused bias in density estimates without making any real difference to the spike at zero. It also did not make much difference to the model fit. It therefore appears that the chital density estimates derived from the forward truncated dataset may be an underestimate, and the estimates from the forward untruncated data may be taken as more accurate. It may have been a better idea to forward truncate the observations in the dry and moist deciduous forests only.

The confidence intervals for density obtained for gaur are relatively wide compared to the other species. This may be a consequence of large variability in cluster size (1 – 36 animals) combined with the low number of detections in some habitats. In some bootstrap runs, most of the resampled temporal replicates may not have any animals detected giving a low density estimate. In others, temporal replicates where large cluster sizes were detected may be selected several times, giving high estimates of density.

Density Estimates- Forward Truncated vs. Untruncated

Generally the forward truncated datasets give a 10 –15 % lower estimate of species densities than the forward untruncated ones. Statistical tests cannot be done to test whether they are significantly different, since they derived from subsets of the same data. The decrease in estimated density is most pronounced in muntjac, which showed an average decrease of 37% when the data is forward truncated. For this species there appears to be so much data on the line that even if a half normal is chosen for the forward untruncated data, as done in this analysis, this may overestimate density. It is also possible, however that the estimates from the forward

truncated data are underestimates. Forward truncation may have removed important observations near the line and thus made it less steep, leading to a lower intercept, though care was taken to avoid try and this. Not all observations at large distances ahead of the observer are actually at zero- some of these are near the line, but forward truncation removes them as well.

Table 4.1 compares the mean point estimates of density obtained for the whole site from this study to those obtained from Karanth & Sunquist, (1992) and Karanth & Sunquist (1995) which were collected from the same transects. Only the point estimates are compared since the confidence intervals from this study are likely to be too narrow, since the spatial variation between different lines in a habitat is not taken into account due to the survey and bootstrap design. Also some of the years in which data was collected overlap between this study and the published ones, so they may not be independent. The estimates from Karanth & Sunquist (1995) may be more accurate than those from Karanth & Sunquist (1992), since the former is based on more data.

Species	Estimated Density (Forward Untruncated)	Estimated Density (Forward Truncated)	Estimated Density (Karanth & Sunquist 1992)	Estimated Density (Karanth & Sunquist 1995)
muntjac	4.3	2.7	4.2	4.3
Wild Pig	4.1	3.5	4.2	3.4
Chital	47.1	41.4	50.6	49.1
Sambar	3.1	2.6	5.5	3.4
Elephant	2.1	1.9	3.3	NA
Gaur	6.6		9.6	5.6

Table 4.1: Comparison of mean densities for entire site for forward untruncated and forward truncated data with estimates from Karanth & Sunquist (1992 and 1995)

The density estimates for the muntjac from the forward untruncated data are very similar to those from both the studies compared. The density from the forward truncated data, though, is much lower. The densities for the pig are not extremely different, and neither are those for the chital when considering that the estimate from

the forward truncated data for chital is likely to be an underestimate. Sambar densities both for the forward truncated and untruncated cases, appear to be lower than that estimated from the other two studies, though the estimates from the forward untruncated data and from Karanth & Sunquist (1995) are quite similar. Karanth & Sunquist (1992) also noted that their estimate of elephant density was likely to be an overestimate, and was probably close to 1.9 animals /sq.km, which tallies well with the results from both the forward truncated and untruncated data.

It therefore appears that the densities for the sambar may have been overestimated in previous studies, possibly due to small sample sizes or the issue of the spike at zero, if it was present at that time also. Muntjac estimates may be much lower than expected if one is willing to accept the validity of such severe forward truncation in this species. These results, however, will need to be seen in the context of field observations, since there could be several factors that this analysis does not take into consideration. Those with more knowledge and experience of the study site may therefore choose which of the density estimates is more likely to reflect reality.

The ranking of species by mean abundance over time gives similar results for the forward truncated and untruncated datasets- chital is the most abundant by a wide margin, and sambar and elephant are the least abundant. The densities of pig and muntjac are quite similar, though the differences are larger for the forward truncated data. In general, they are similar to the rankings derived from Karanth & Sunquist (1992 and 1995) and Madhusudan & Karanth (2000).

4.2 Trends

The results from the trend analysis suggest that there is very overall little trend in animal densities; the populations are generally stable. In most cases, the smooth terms for the years were not significant. Trends for the forward truncated and forward untruncated data were also very similar. This is to be expected, given that this is a relatively short time series and Nalkeri is a well protected and managed area (Madhusudan & Karanth 2000). Animal populations may therefore be close to the carrying capacity.

For sambar, trends in the dry deciduous stratum were found to be highly significant for both the forward truncated and untruncated data. The GAM suggests that the population is growing at an exponential rate –this may be an artefact of the relatively small number of years. The associated 95% confidence intervals for the degrees of freedom are 1 and 1.35 indicating that these trends are more linear than smooth. There also seems to be some indications of a similar upward trend in the gaur and pig in dry deciduous forests as well, though they are not significant. It would be interesting to see if the densities for later years follow this trend.

muntjac and chital were the other species where there was a significant trend, again in the dry deciduous forest. The degrees of freedom fit for muntjac for both the forward truncated and untruncated datasets (2.0- 2.9) are relatively high compared to the other species, indicating smooth trends. The upper confidence limit of the degrees of freedom is close to the maximum number (3) allowed for the smooth, and so this should possibly be increased. The effect of forward truncation is apparent as well - the curves for the forward truncated dataset are at a lower level, with the peaks cut down. The densities seem to increase until around 1995, following which there is a downturn and the population returns to its previous state.

The number of degrees of freedom fit for chital in dry deciduous forest were closer to 1, so a linear term may be better. The trend on first glance appears logistic, though in reality it is probably similar to the trends in the muntjac. This is for the forward truncated dataset, which may not be very accurate; however, considering that the trends in other species are similar for both forward truncated and untruncated datasets, this is likely to hold true for the untruncated chital dataset as well.

It is interesting that all the significant trends took place in the dry deciduous stratum. This habitat has generally low species densities compared to the more moist regions (Karanth & Sunquist 1992; Karanth et al 1999). These increases and decreases could have something to do with water availability.

4.3 Improvements, Recommendations and Extensions

Survey Design & Execution

The design of the transects, especially in thickly forested areas, contributed to the spike at zero. Density estimates may be subject to some positive bias in the presence of this spike. This problem is difficult to correct at the analytical stage; forward truncation is essentially a subjective process and may lead to bias as well. Model fit can also be improved for data without a spike, especially when fitting covariates. Solutions to this include:

- Redesigning transects so that they are not completely straight, but are made up of zig-zag lines. This should make the analysis less vulnerable to unmodelled animal movement ahead of the observer. The difference in angle between legs of the transect need not be very high, perhaps around 20°- it needs to be just enough so that observers do not have very long views along

the line. Each leg may be around 50 metres long, since most observations beyond this distance ahead of the observer tend to be recorded as on the line. This may also have the benefit of cutting down on cases of sighting angles being rounded to zero, since animals far ahead of the observer in the vicinity of the line have a greater chance of being recorded as on the line. A disadvantage of this could be that during surveys, an observer appearing suddenly 'around the corner' may surprise animals. This may be dangerous if the animal in question is, for example, an elephant.

- An alternative would be to train observers not to record animals that are at very large distances ahead of them. This could be difficult to implement since it may not always be easy to judge when to and when not to record an animal far ahead of the observer, especially for small species like the muntjac.
- If only even degrees are marked on compasses, they may be changed so that all degrees are marked to avoid rounding angles to the nearest even number. Observers also need to overcome the natural instinct to record even angles, and especially must try to avoid rounding sighting angles to zero.
- The purposive placement of these transects may lead to bias in density estimates, since the few transects may not take into consideration spatial variation in animal densities within habitat strata. The estimates from this analysis may be subject to such a bias. A new survey design with much more spatial replication has in fact been implemented and is in use for the past 2-3 years, which should give more accurate results.

Analysis

Forward truncation was essentially a subjective process, and therefore prone to bias. Its validity will need to be further established, perhaps through more quantitative criteria. Another issue is that the density estimates for the forward untruncated data were obtained through multiple covariate distance sampling. This may have biased the detection function due to its vulnerability to a spike in the data when fitting covariates (section 2.1). Repeating the analysis for the untruncated data with conventional distance sampling models would shed more light on this issue. Other improvements to the analysis include:

- Forward truncation for chital only in forested habitats
- Fitting separate detection functions by habitat type and deriving density estimates from these for chital since the sample size is very large.
- A gamma errors model may have led to different density estimates
- Increasing the number of bootstrap resamples to around 1000, which could not be done in this study due to time constraints would give more accurate confidence intervals
- Incorporating model selection uncertainty in the bootstrap, by making the program select between different models (such as whether to fit separate smooths for different habitats, or a global smooth) in each resample. In species where there is not much difference in trends between habitats, GAMs could be fitted at the global level rather than for each habitat.

Extensions

Possible extensions to this project include:

- Spatial variables could be included in the GAM analyses, which would give a spatio-temporal model. An analysis on this was begun, and preliminary results

showed that latitude and longitude were highly significant. The model was however a poor one due to the relatively large number of zeros; a zero inflated model could possibly be useful in this. Unfortunately this could not be pursued for lack of time.

- The trends in herbivore density could be compared with trends in carnivore densities to see if there is any correlation between these.
- The herbivore densities could be used to predict tiger densities, as in Karanth et al (2004).
- Over the long term, other regions of the park may be monitored to give a comprehensive spatio-temporal model of animal densities in Nagarhole.

Conclusion

The results from this study indicate that density estimates from highly spiked datasets may overestimate populations. However, considering that two major assumptions of distance sampling- no animal movement and the shape criterion (Buckland et al 2001)- appear to be violated in this dataset, the results from the forward untruncated dataset do not seem extremely poor. This problem, however, will need to be taken into account in the future.

Animal population trends in Nalkeri seem stable, which is probably due to the effective protection offered to them. Some of the population trajectories, however, seem to be slightly downwards - they are not significant, but it may be worth keeping a watch on them in the future.

This study contributed to the need for information on long term animal population trends in India, and it serves as a basis for further research. The results

from this study could eventually be incorporated into a large scale model of herbivore-carnivore population dynamics, or a spatio-temporal model at Nagarhole.

REFERENCES

Borchers, D.L., Buckland, S.T. and Zucchini, W. (2002). Estimating Animal Abundance. Springer-Verlag, London.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L. (2001). Introduction to Distance Sampling. Oxford University Press, New York.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L. (2004). Advanced Distance Sampling. Oxford University Press, New York.

Fewster, R.M., Buckland, S.T., Siriwardena, G.M, Baillie, S.R and Wilson, J.D (2000). Ecology 81(7) :1970-1984.

Fox, J. (2002). An R and S-Plus companion to applied regression. Sage Publications USA.

Hastie, T.J. and Tibshirani, R.J. (1990). Generalized Additive Models. Chapman & Hall, London.

Jathanna, D., Karanth, K.U. and Johnsingh, A.J.T (2003). Estimation of large herbivore densities in the tropical forests of southern India using distance sampling. Journal of the Zoological Society of London 261 :285-290.

Karanth, K. U. & Sunquist, M. E. (1992). Population structure, density and biomass of large herbivores in the tropical forests of Nagarhole, India. Journal of Tropical Ecology 8: 21±35.

Karanth, K.U. and Sunquist, M. E. (1992). Population structure, density and biomass of large herbivores in the tropical forests of Nagarhole, India. Journal of Tropical Ecology 8:21-35.

Karanth, K.U. and Sunquist, M. E. (1995). Prey selection by tiger, leopard and dhole in tropical forests. Journal of Animal Ecology 64:439-450.

Karanth, K.U. and Nichols, J.D. (1998). Estimation of tiger densities in India using photographic captures and recaptures. Ecology 79(8): 2852-2862.

Karanth, K.U. and Sunquist, M. E. (2000). Behavioural correlates of predation by tiger (*Panthera tigris*), leopard (*Panthera pardus*) and dhole (*Cuon alpinus*) in Nagarhole, India. Journal of the Zoological Society of London 250:255-265.

Karanth, K.U. and Stith, B.M. (1999). Prey depletion as a critical determinant of tiger population viability. Pages 100-113 in Riding the Tiger: : Tiger conservation in human-dominated landscapes. (editors: Seidenstickers, J., Christie, S. and Jackson, P.). Cambridge University Press, Cambridge.

Karanth, K.U., Sunquist, M. E. and Chinnappa, K.M. (1999). Long-term monitoring of tigers: lessons from Nagarhole. Pages 114-122 in Riding the Tiger: : Tiger conservation in human-dominated landscapes. (editors: Seidenstickers, J., Christie, S. and Jackson, P.). Cambridge University Press, Cambridge.

Karanth, K.U., Nichols, J.D., Kumar, N.S., Link, W.A. and Hines, J.E. (2004). Tigers and their prey: predicting carnivore densities from prey abundance. *Proceedings of the National Academy of Sciences* 101(14) 4854-4858.

Madhusudan, M.D., Karanth, K.U. (2000). Hunting for an answer: is local hunting compatible with large mammal conservation in India? Pages 339-354 in *Hunting for Sustainability in Tropical Forests* (editors: Robinson, J.G. and Bennett, E.L.). Columbia University Press, USA.

Kumar, N.S. (2000). Ungulate density and biomass in the tropical semi-arid forest of Ranthambore, India. MSc. Thesis, Pondicherry University.

Karanth, K.U. and Nichols, J.D. (editors) 2002. *Monitoring tigers and their prey- a manual for researchers, managers and conservationists in tropical Asia*. Centre for Wildlife Studies, India.

Marques, F.F.C., and Buckland, S.T. (2004). Covariate models for the detection function. In *Advanced Distance Sampling* (Editors: Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L.). Oxford University Press, New York.

Prater, S.H. (1965). *The Book of Indian Mammals*. Bombay Natural History Society, Bombay.

R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Ramakrishnan, U., Coss, R.G. and Pelkey, N.W. (1998). Tiger decline caused by the reduction of large ungulate prey: evidence from a study of leopard diets in southern India. *Biological Conservation* 89: 113-120.

Sunquist, M.E, Karanth, K.U and Sunquist, F. (1999). Ecology, behaviour and resilience of the tiger and its conservation needs. Pages 5-17 in *Riding the Tiger: : Tiger conservation in human-dominated landscapes*. (editors: Seidenstickers, J., Christie, S. and Jackson, P.). Cambridge University Press, Cambridge.

Thomas, L., Burnham, K.P. and Buckland, S.T. (2004). Temporal inferences from distance sampling surveys. In *Advanced Distance Sampling* (Editors: Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L.). Oxford University Press, New York.

Thomas, L. (2001). *WCS- India Ungulate Density Project Consultancy Report*. Unpublished.

Thomas, L., Laake, J.L., Strindberg, S., Marques, F.F.C., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Hedley, S.L., Pollard, J.H., Bishop, J.R.B. and Marques, T.A. 2005. *Distance 5.0 Beta 4*. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. <http://www.ruwpa.stand.ac.uk/distance/>

Thomas, L. (1996). Monitoring long-term population change: why are there so many analysis methods? *Ecology* 77 (1): 49-58

Venables, W.N. and Ripley, B.D. (2002). *Modern Applied Statistics with S* (4th ed). Springer-Verlag, New York.

Wickramanayake, E.D., Dinerstein, E., Robinson, J.G., Karanth, K.U., Rabinowitz, A., Olson, D., Mathew, T., Hedao, P., Conner, M., Hemley, G. and Bolze, D. (1998). An ecology based method for defining priorities for large mammal conservation: the tiger as a case study. *Conservation Biology* 12(4): 865-878.

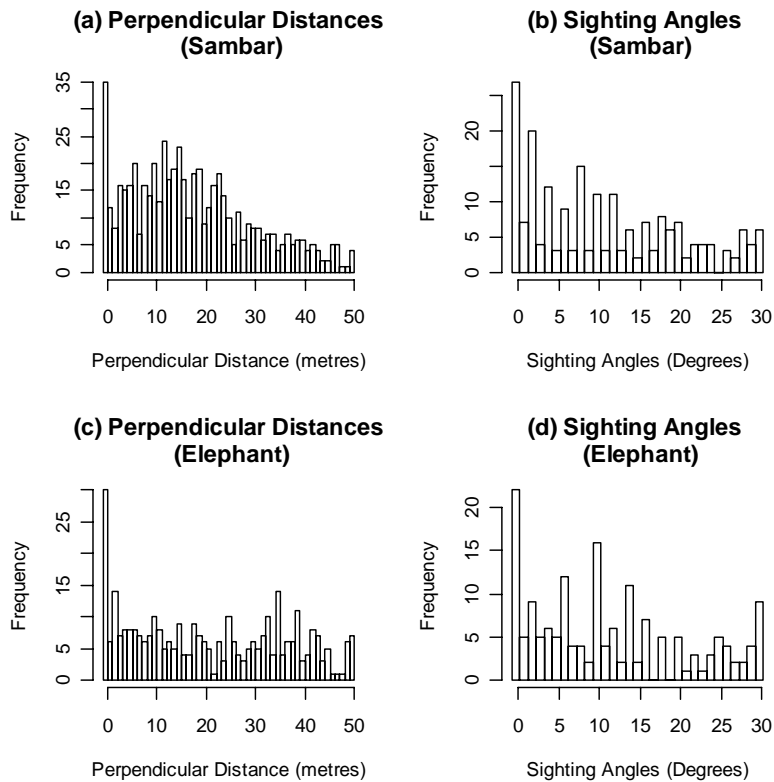
Wood, 2001. mgcv: GAMs and Generalized Ridge Regression for R. *R News* vol 1/2.

Wood, S.N. (2000) Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties. *Journal of the Royal Statistical Society* 62(2):413-428.

APPENDICES

Appendix 1

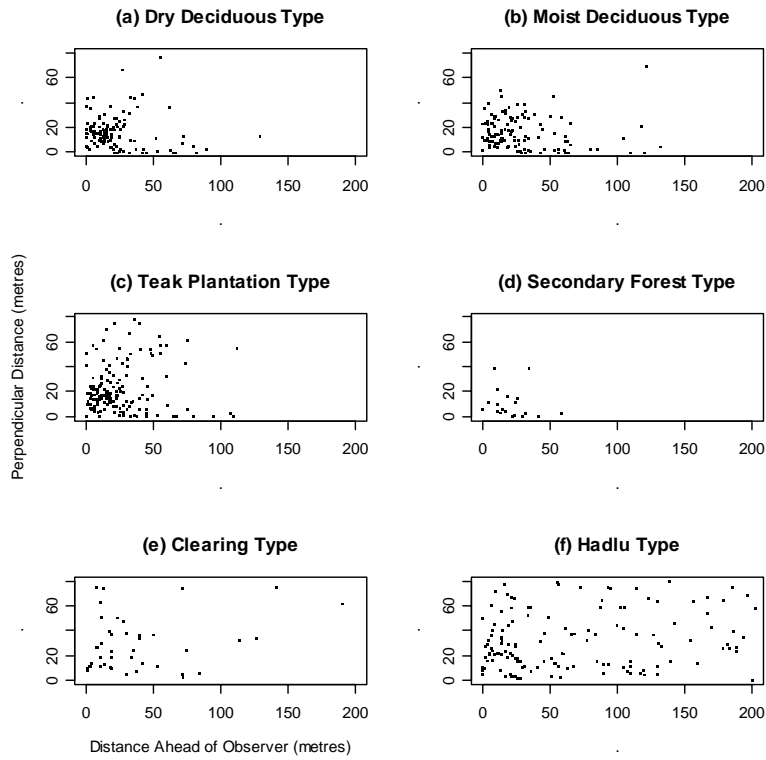
Forward Truncation for Sambar and Elephant



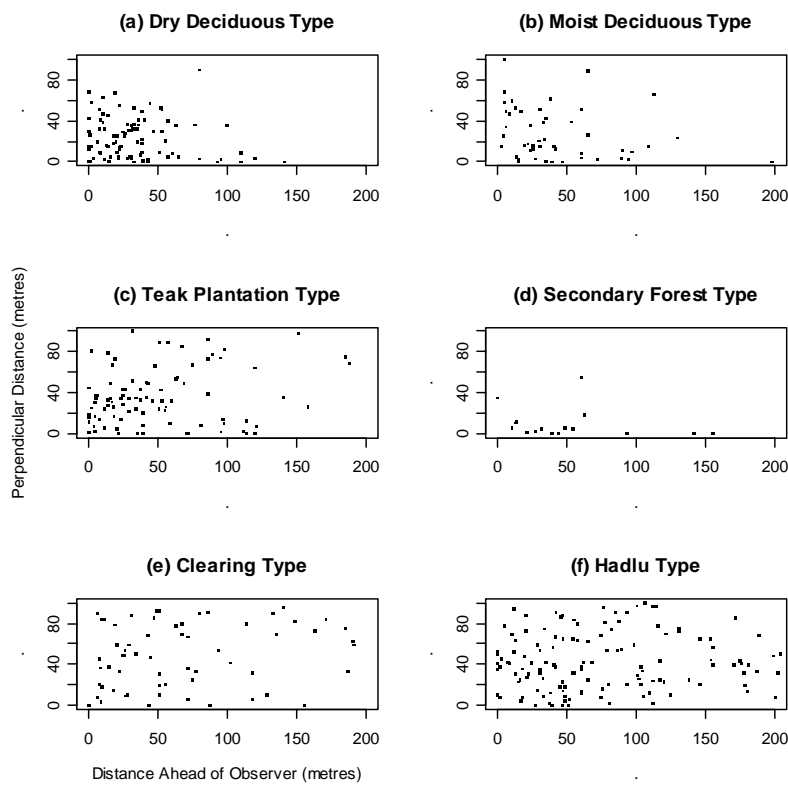
The datasets of Sambar and Elephants have certain issues in common. For both species it appears that the histogram in the first 20 metres is not unduly steep (figure 5.1a and 5.1c); however, there is a definite spike at zero. Again, even sighting angles are favoured, particularly for sambar (figures 5.1 b and 5.1d).

Distance plots by habitat type for sambar (figure 5.2 a-f) and elephant (figure 5.3 a-f) reveal differences between the forested and open habitat types- there appears to be a correlation between observations at zero and large distances ahead of the observer, especially for sambar. The plots for the open habitats, on the other hand, do not show such patterns.

It was found that the transect which elephants were suspected to be using as trails did not contribute much to the zeros. This possibility was therefore ignored.



plots for sambar



plots for elephant

For both these species, therefore, it appears that straight lines as well as rounding of sighting angles contribute to the spike, with the former being particularly

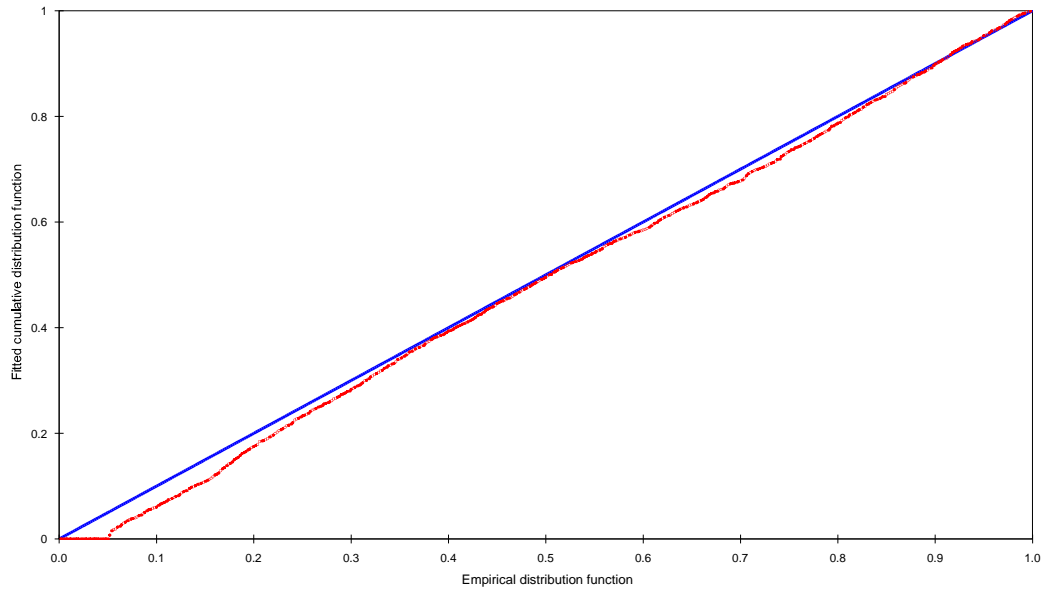
noticeable in the forested habitat types. It is also apparent that forward truncation at the level of the entire site would lead to a much loss of data from the hadlus.

It was therefore decided to carry out forward truncation of observations in the forested habitat types (dry deciduous, moist deciduous, teak plantations and secondary forest).

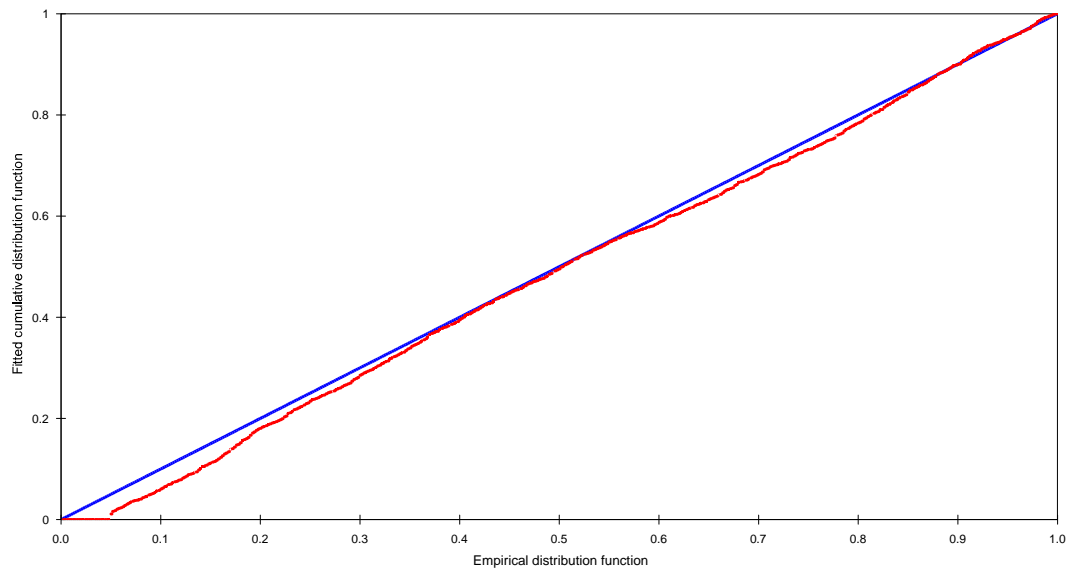
Appendix 2 Quantile-Quantile Plots by Species

The following graphs compare the Q-Q plots obtained for the forward truncated and the forward untruncated datasets for each species.

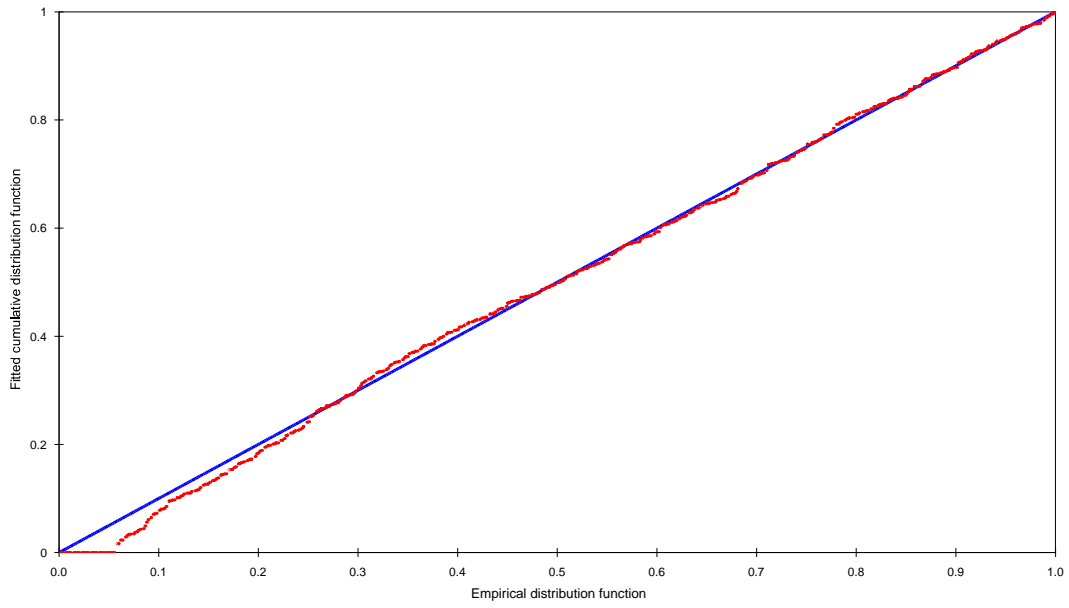
Chital (Untruncated)



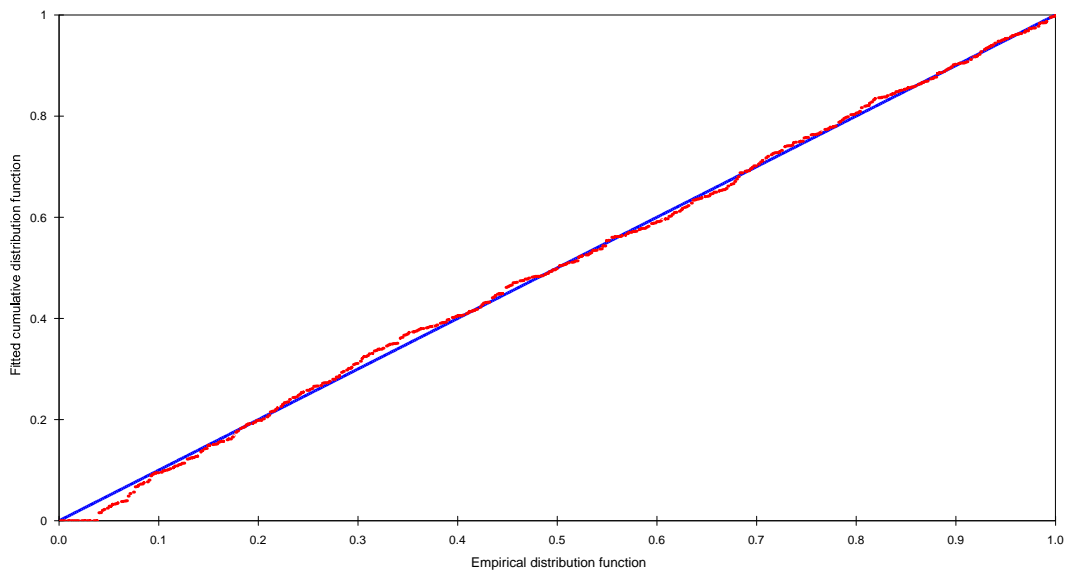
Chital (truncated)



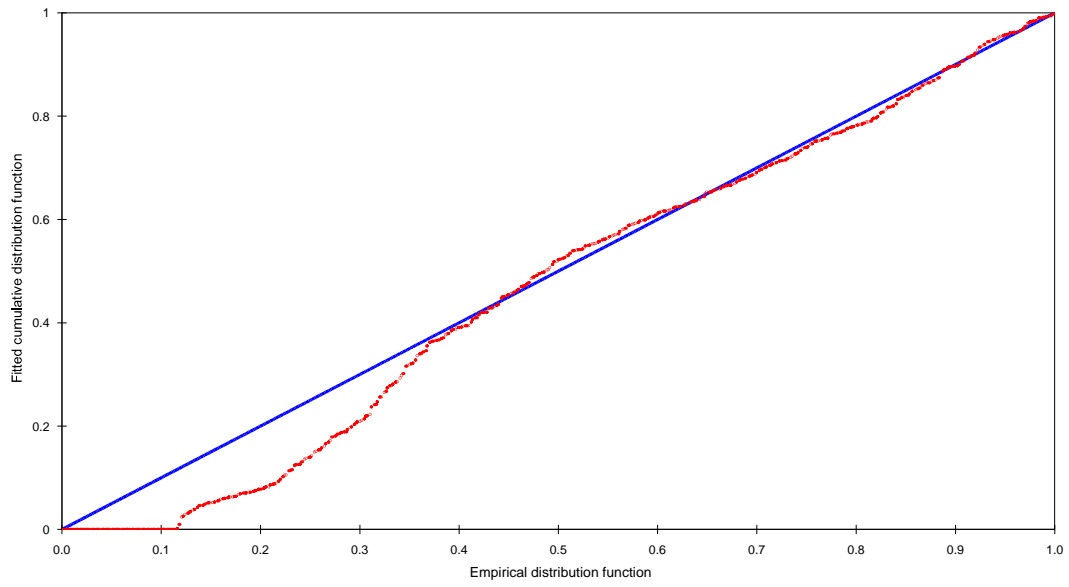
Sambar (untruncated)



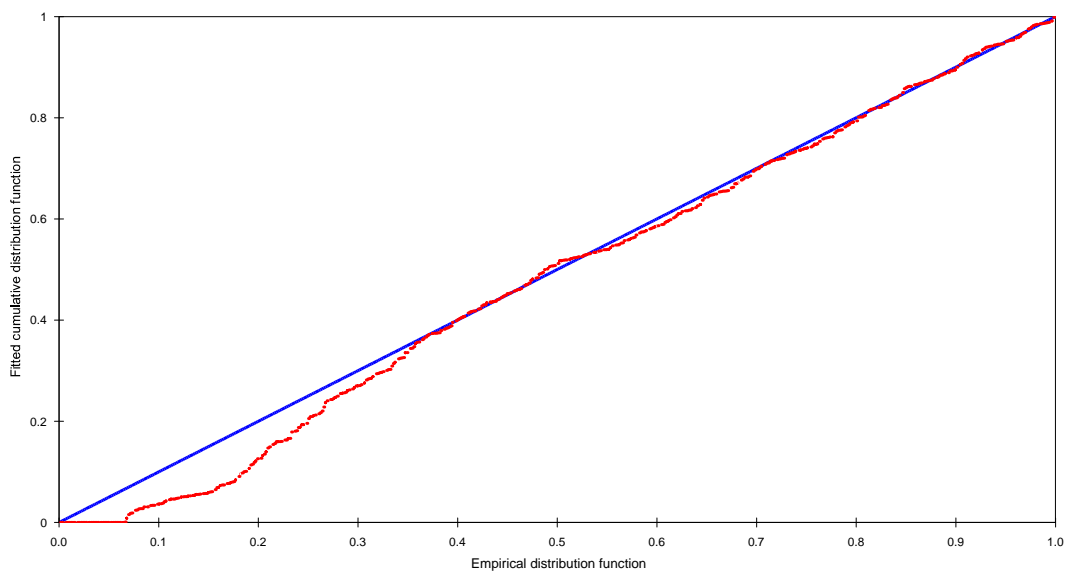
Sambar (truncated)



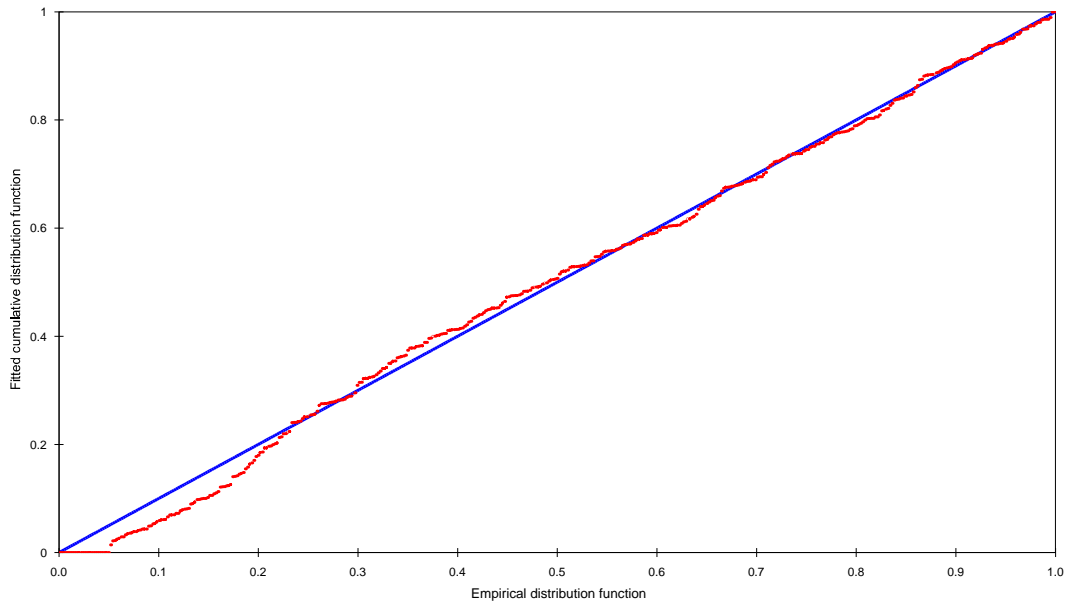
muntjac (untruncated)



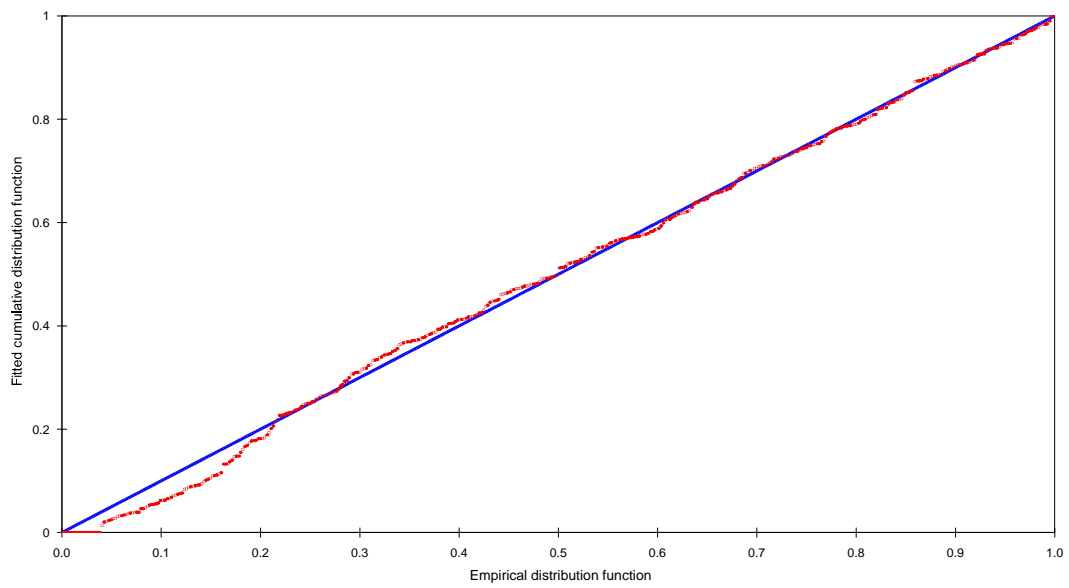
muntjac(truncated)



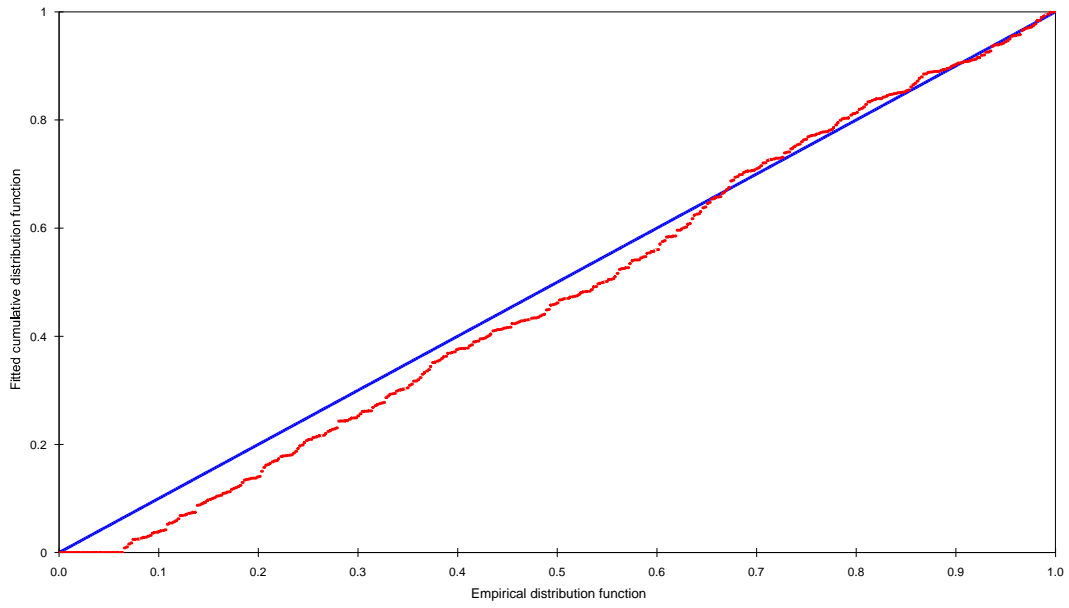
Wild Pig (untruncated)



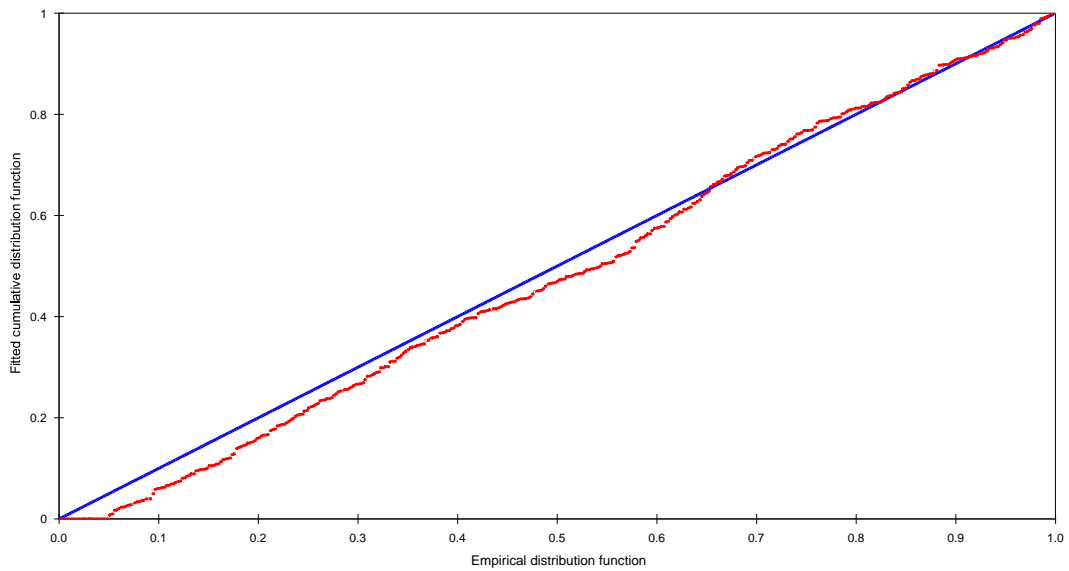
Wild Pig (truncated)



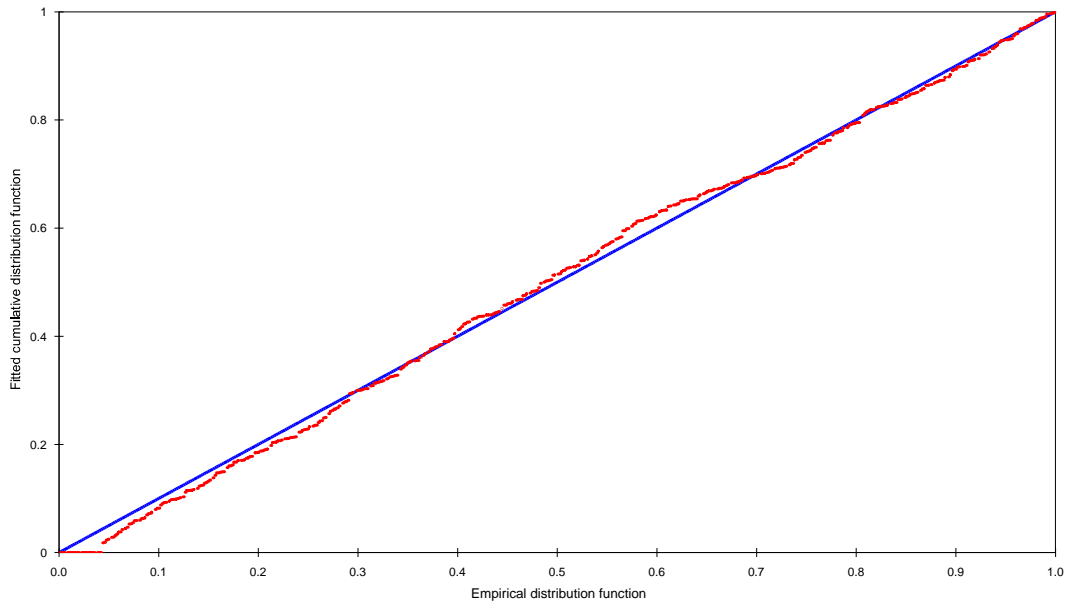
Elephant (untruncated)



Elephant (truncated)



Gaur



Appendix 3

Density estimates over time for the entire site as well as by habitat stratum for all species

Overall

Wild Pig

	Forward Untruncated	4.084946807	Forward truncated	3.544993373
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	4.93	2.5 - 7.1	4.14	2.5 - 6.3
1990	6.07	3.6 - 8.3	4.97	2.9 - 6.8
1991	4.93	3 - 7.7	4.30	2.4 - 6.7
1994	2.32	1.3 - 3.2	2.16	1.4 - 3.2
1995	3.90	2.8 - 5.2	3.11	2 - 4.3
1996	5.13	3.2 - 7.4	4.75	2.9 - 7.5
1997	2.13	0.93 - 3.5	1.76	0.8 - 2.7
1998	2.72	1.5 - 4.1	2.51	1.3 - 3.7
1999	3.52	2 - 5.2	3.19	2 - 4.7
2000	5.21	2 - 7.8	4.57	2.4 - 7.5

Chital

	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	55.00	NA	40.42	33.5 - 47.3
1990	45.63	NA	37.29	29.4 - 44.6
1991	42.82	NA	37.89	29.8 - 43.8
1994	42.39	NA	39.70	34.3 - 46.3
1995	28.82	NA	32.04	25.8 - 37
1996	54.13	NA	42.59	35 - 49.3
1997	59.70	NA	50.94	43.1 - 59.4
1998	45.29	NA	43.86	38.5 - 50.1
1999	50.76	NA	48.05	37.7 - 59.5
2000	46.61	NA	41.15	35.6 - 47.5

Sambar

	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	2.67	1.7 - 4	2.21	1.24 - 3.8
1990	3.43	2.3 - 4.9	2.72	1.6 - 4.3
1991	3.34	2.1 - 4.4	2.89	1.8 - 3.9
1994	3.43	2.2 - 4.8	2.91	2 - 4.7
1995	2.87	1.8 - 4	2.37	1.4 - 3.5
1996	3.11	2.1 - 4.5	2.61	1.7 - 3.9
1997	2.20	1.3 - 2.9	1.99	1 - 2.8
1998	2.95	2 - 4.1	2.56	1.8 - 3.8
1999	3.36	2.2 - 4.8	2.80	1.9 - 4.1
2000	3.28	2.2 - 4.5	2.80	1.9 - 4.1

Elephant

Year	Forward Untruncated		Forward truncated	
	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	3.01	1.7 - 5.8	2.75	1.6 - 4.8
1990	1.71	0.77 - 3.3	1.65	0.73 - 3
1991	2.44	1.6 - 4.1	2.22	1.3 - 3.6
1994	2.90	1.7 - 4.4	2.62	1.7 - 4
1995	0.78	0.37 - 1.3	0.64	0.37 - 0.99
1996	0.43	0.23 - 0.74	0.43	0.21 - 0.75
1997	1.78	1.2 - 3.1	1.77	1.2 - 2.9
1998	2.95	2.0 - 4.6	2.74	2 - 3.9
1999	1.75	1.1 - 3.0	1.61	0.86 - 2.7
2000	3.24	2.1 - 4.9	3.00	2.1 - 4.7

Gaur

Year	Forward Untruncated			
	Density (/sq km)	95% Confidence intervals		
1989	6.43	3.3 - 12.5		
1990	6.01	3.2 - 11.8		
1991	8.18	4.9 - 13.1		
1994	4.96	2.4 - 9.8		
1995	4.81	2.9 - 7.7		
1996	7.15	4.7 - 11.7		
1997	5.57	3.5 - 10		
1998	6.90	4.2 - 12.6		
1999	10.32	6.4 - 19.3		
2000	5.85	3.2 - 11.3		

*Dry Deciduous Stratum**muntjac*

Year	Forward Untruncated		Forward Truncated	
	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	0.89	0 - 2	0.462701551	0 - 1.5
1990	1.31	0.53 - 2.1	0.492063667	0.18 - 1
1991	0.99	0 - 2	0.598222311	0 - 1.4
1994	3.53	1.6 - 6.2	1.996197515	1.1 - 3.9
1995	6.62	4.3 - 9.4	3.906172141	3.1 - 5.6
1996	6.50	3.4 - 10	3.943796143	2.1 - 6.7
1997	6.31	3.5 - 9	3.828500708	2.5 - 5.7
1998	7.87	4.5 - 9.2	4.737581229	3.2 - 6.7
1999	3.43	1.5 - 5.9	2.583292785	1.5 - 4.4
2000	1.32	0.43 - 2.4	0.890419872	0.41 - 1.6

Pig

Year	Forward Untruncated		Forward truncated	
	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	0.30	0 - 0.83	0.28	0 - 0.8
1990	0.43	0 - 1.7	0.40	0 - 1.6
1991	1.49	0 - 4.1	1.39	0.14 - 3.8
1994	0.89	0 - 2	0.83	0.12 - 2

1995	0.29	0 – 0.93	0.27	0 - .85
1996	0.74	0 – 1.8	0.69	0 - 1.4
1997	1.86	0 – 5.3	1.74	0 - 4.9
1998	2.70	0.53 - 5.7	2.52	0.52 - 5.7
1999	3.88	1.3 - 6.7	3.62	1.2 - 6.7
2000	1.94	0.15 - 5.1	1.81	0.12 - 4.7
<i>Chital</i>				
	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	3.97	NA	3.61	1.5 - 6.5
1990	2.56	NA	2.51	1 - 4.4
1991	4.77	NA	5.00	1.8 - 7.8
1994	7.92	NA	8.48	5.5 - 11.8
1995	10.61	NA	12.23	8.8 - 16.3
1996	15.40	NA	15.45	8.1 - 22
1997	9.65	NA	9.68	5.7 - 13.5
1998	9.73	NA	10.33	6.5 - 15
1999	12.89	NA	14.77	9.6 - 21.5
2000	9.34	NA	10.13	7.3 - 13.4

Sambar

	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	0.12	0 – 0.44	0.10	0 - 0.34
1990	1.15	0.1 - 2.5	0.88	0 - 2.4
1991	0.81	0 – 0.91	0.54	0 - 0.34
1994	1.38	0.61 - 2.4	0.98	0.39 - 1.8
1995	1.89	0.88 - 3.2	1.48	0.57 - 2.5
1996	0.81	0.26 - 1.4	0.62	0.19 - 1.2
1997	0.89	0 - 0.94	0.71	0 - 0.58
1998	3.24	1.7 - 5	2.76	1.5 - 4.6
1999	6.44	3.5 - 9.8	5.25	2.8 - 8.5
2000	3.00	1.5 - 4.5	2.75	1.5 - 4.6

Elephant

	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	0.95	0 – 2.3	0.89	0.11 - 2.1
1990	1.60	0 – 6.4	1.44	0 - 5.2
1991	3.24	1.4 - 6.7	2.81	0.91 - 5.6
1994	3.81	1.8 - 6.8	3.42	1.5 - 6.3
1995	0.88	0.05 - 2.4	0.40	0.06 - 1.1
1996	0.00	0.00	0.00	0.00
1997	1.62	0.47 - 3.4	1.58	0.49 - 3.7
1998	2.54	1 – 5.5	2.17	0.84 - 4.5
1999	1.15	0.37 - 2.4	1.02	0.25 - 2.1
2000	1.54	0.60 - 3.3	1.40	0.54 - 2.9

Gaur

	Forward Untruncated			
--	---------------------	--	--	--

Year	Density (/sq km)	95% Confidence intervals		
1989	1.44	0.0 - 6.6		
1990	5.00	0.51 - 15.5		
1991	4.93	0.83 - 15.1		
1994	2.55	0.44 - 8.3		
1995	0.96	0.0 - 3.1		
1996	0.78	0.09 - 2.35		
1997	6.34	2.6 - 12.7		
1998	3.07	0.8 - 7.8		
1999	7.24	3.3 - 17.6		
2000	6.54	2.5 - 14.9		

Moist Deciduous stratum

<i>muntjac</i>				
	Forward Untruncated		Forward Truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	6.233695321	3.4 - 9.1	3.67685824	1.8 - 6.2
1990	5.241926675	2.6 - 9	2.247248586	1.1 - 4.2
1991	3.988790479	2 - 6.4	2.732469146	1.7 - 5
1994	6.96241149	3.3 - 12.9	3.237588783	1.5 - 5.1
1995	9.312860474	5.4 - 13.2	6.109964423	3.8 - 9.7
1996	13.23331965	8.4 - 20	9.728788578	5.9 - 16
1997	3.339197784	1.5 - 5.4	2.206712077	0.97 - 3.8
1998	7.563927466	4.6 - 10.8	4.354751884	2.4 - 7
1999	3.254041261	1.5 - 5.6	2.437201369	0.85 - 4.5
2000	6.184631028	3.5 - 9.1	4.222474872	2.6 - 7.3
<i>Wild Pig</i>				
	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	9.16	3.6 - 15.8	6.99	3 - 12.8
1990	11.31	5.4 - 16.9	8.25	3.7 - 13
1991	10.49	4.1 - 17.4	8.74	3.9 - 14.5
1994	0.79	0 - 2	0.71	0 - 1.5
1995	7.96	4.8 - 12.1	5.67	2.8 - 8.6
1996	8.46	3.9 - 14	7.67	3.9 - 13.3
1997	2.61	0.52 - 4.9	1.68	.33 - 3.2
1998	3.86	1.4 - 7.2	3.49	1.2 - 6.5
1999	5.29	1.5 - 9.5	4.63	1.8 - 9.1
2000	13.87	5.4 - 21.6	11.97	5.1 - 21
<i>Chital</i>				
	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	38.79	NA	34.52	20.6 - 49.5
1990	29.65	NA	26.72	20.1 - 33.6
1991	28.94	NA	25.47	16.1 - 33.5
1994	28.44	NA	26.98	15.3 - 39.9
1995	20.43	NA	21.25	14.7 - 27.6
1996	28.58	NA	26.78	16.6 - 34.1
1997	25.90	NA	21.03	13.7 - 27.6

1998	21.25	NA	22.47	17.3 - 29.2
1999	23.45	NA	24.61	14.4 - 33.6
2000	31.60	NA	30.59	22.8 - 38.2
<i>Sambar</i>				
	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	2.16	0.82 - 3.7	1.84	0.57 - 3.8
1990	5.31	2.8 - 8.6	4.05	1.9 - 7.3
1991	2.09	1.1 - 3.6	1.78	0.68 - 3
1994	1.57	0.28 - 3.3	1.40	0.21 - 3.3
1995	2.79	2.8 - 4.8	2.07	0.55 - 3.9
1996	1.65	0.74 - 2.7	1.07	0.36 - 2.01
1997	1.13	0.63 - 1.9	1.03	0.34 - 2.4
1998	1.66	0.7 - 3.1	1.27	0.46 - 2.6
1999	0.65	0.21 - 1.3	0.57	0.12 - 1.2
2000	3.11	1.3 - 5.3	2.63	1.1 - 4.5
<i>Elephant</i>				
	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	0.89	0 - 2.5	0.88	0 - 2.5
1990	1.15	0.20 - 2.1	1.15	0.24 - 2.4
1991	1.13	0.19 - 2.5	0.91	0.08 - 2.2
1994	0.09	0 - 0.26	0.10	0 - 0.29
1995	0.34	0 - 0.83	0.35	0 - 0.76
1996	0.13	0 - 0.4	0.12	0 - 0.33
1997	0.44	0.05 - 1.3	0.44	0.05 - 1.22
1998	0.51	0.06 - 1.3	0.37	0.03 - 1.1
1999	1.97	0.65 - 4.5	1.91	0.53 - 3.9
2000	0.87	0.24 - 1.8	0.93	0.14 - 1.8
<i>Gaur</i>				
	Forward Untruncated			
Year	Density (/sq km)	95% Confidence intervals		
1989	1.59	0.11 - 4.1		
1990	6.27	0.2 - 17.7		
1991	1.95	0.0 - 6.7		
1994	6.18	0.67 - 18.4		
1995	1.09	0.0 - 3.4		
1996	4.22	0.75 - 12.3		
1997	4.19	1.6 - 11.4		
1998	10.68	4.1 - 26.4		
1999	12.48	4.8 - 35.8		
2000	6.94	2.1 - 19.2		

Teak Dominant Stratum

<i>munjtjac</i>				
	Forward Untruncated		Forward Truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals

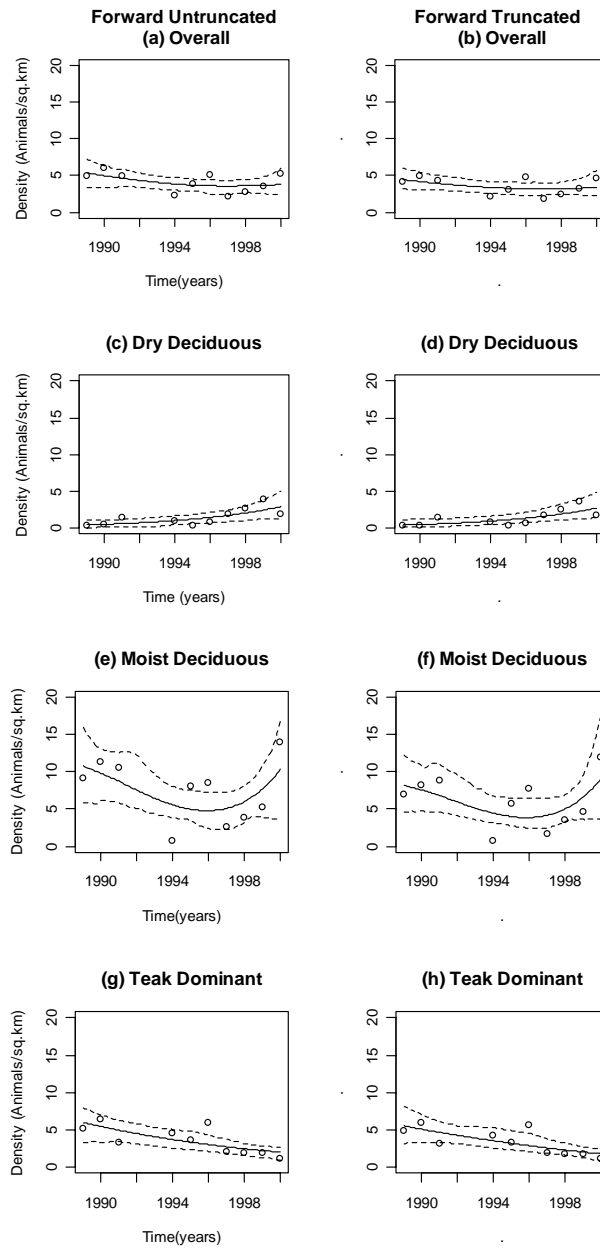
1989	5.017529729	3.3 - 7.1	3.930898551	2.6 - 6
1990	3.819899638	2.4 - 5.8	1.737898176	1.1 - 2.8
1991	3.735244293	2.1 - 6.5	2.198099943	1.5 - 3.8
1994	3.946769564	2.2 - 5.9	2.494273208	1.5 - 4
1995	4.620838001	2.8 - 7.2	2.838719869	1.8 - 5.1
1996	2.568149337	1.1 - 4.2	1.253206154	0.3 - 2.5
1997	1.60397619	0.7 - 2.7	1.038338028	0.52 - 1.9
1998	1.249547548	0.44 - 2.2	0.853905406	0.29 - 1.6
1999	1.407122643	0.29 - 2.9	1.33737306	0.43 - 2.5
2000	1.985205124	0.9 - 3.5	1.304628111	0.7 - 2.4
<i>Wild Pig</i>				
	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	5.221380047	2.1 - 8.2	4.894994913	2.2 - 4.3
1990	6.359574265	2.4 - 10.1	5.931074223	3.1 - 8.8
1991	3.331222276	1.7 - 5.6	3.150066602	1.6 - 6.6
1994	4.544986553	2.2 - 6.5	4.24762035	2.5 - 6.3
1995	3.570919672	1.8 - 5.5	3.331716688	1.8 - 4.8
1996	5.932525803	2.5 - 10.1	5.591048046	2.6 - 11.4
1997	1.9731296	0.61 - 3.4	1.841516481	0.69 - 3.1
1998	1.870893913	0.74 - 3.5	1.754643824	0.64 - 3.3
1999	1.918297878	0.68 - 3	1.787875877	0.65 - 2.9
2000	1.162907164	0.34 - 2.2	1.084129972	0.3 - 2.1
<i>Chital</i>				
	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	105.4445559	NA	72.46402864	59.6 - 86
1990	89.92753834	NA	71.30138013	51.7 - 91.3
1991	81.76095151	NA	71.86673016	53.4 - 81.8
1994	78.70573518	NA	72.65731066	60.6 - 82.6
1995	48.76964724	NA	54.97719327	41 - 64.9
1996	102.3316153	NA	74.81068407	58.3 - 92.6
1997	122.5955068	NA	104.3185659	86.8 - 123.1
1998	89.9824723	NA	85.04934403	73.9 - 97.6
1999	99.63670298	NA	90.58203482	68.1 - 117.4
2000	85.82504153	NA	72.32611861	61.1 - 87
<i>Sambar</i>				
	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	4.970271774	2.8 - 7.9	4.08	2 - 7.6
1990	3.718265149	2.1 - 6	3.11	1.5 - 5.4
1991	6.182033182	3.6 - 9.1	5.48	3.2 - 8.3
1994	6.351764386	3.7 - 10	5.49	3.2 - 8.6
1995	3.66367329	2 - 5.8	3.27	1.7 - 5.5
1996	5.939635933	3.8 - 9	5.26	3.1 - 7.8
1997	3.986418071	2.1 - 5.9	3.66	1.8 - 6
1998	3.690980425	2.3 - 5.9	3.37	2.1 - 5.5
1999	3.075018397	1.8 - 4.5	2.63	1.7 - 3.9
2000	3.611152007	2.2 - 4.5	2.97	1.9 - 4.8

<i>Elephant</i>				
	Forward Untruncated		Forward truncated	
Year	Density (/sq km)	95% Confidence intervals	Density (/sq km)	95% Confidence intervals
1989	6.142193622	3.3 - 12.5	5.55	3 - 10.3
1990	2.202368488	0.88 - 4.4	2.17	0.94 - 3.8
1991	2.835335586	1.3 - 5.1	2.75	1.2 - 5.4
1994	4.324350474	2.1 - 6.8	3.90	2.1 - 6.3
1995	1.042621506	0.46 - 1.9	1.04	0.49 - 1.8
1996	0.988075693	0.45 - 1.7	0.99	0.45 - 1.8
1997	2.913850284	1.8 - 5.2	2.90	1.8 - 4.6
1998	5.082064112	3.6 - 7.5	4.94	3.4 - 7.1
1999	2.035725242	0.94 - 4.0	1.83	0.67 - 3.6
2000	6.283299243	4.1 - 10.4	5.75	3.8 - 9.1
<i>Gaur</i>				
	Forward Untruncated			
Year	Density (/sq km)	95% Confidence intervals		
1989	13.80	5.8 - 28.7		
1990	6.57	3.1 - 15.2		
1991	15.29	9.3 - 23.4		
1994	5.85	2.6 - 10.1		
1995	10.49	6.1 - 17		
1996	14.12	8.5 - 20.9		
1997	6.02	3.0 - 9.3		
1998	6.95	3.2 - 10.8		
1999	11.00	6.1 - 18		
2000	4.51	1.8 - 7.5		

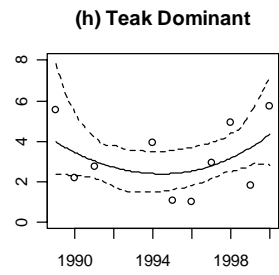
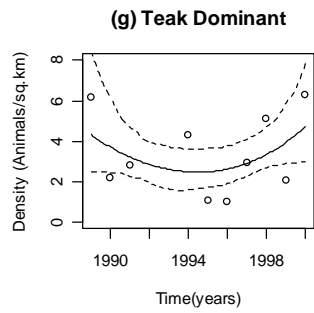
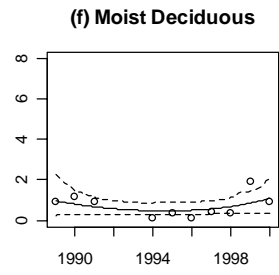
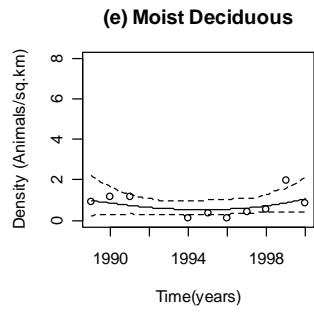
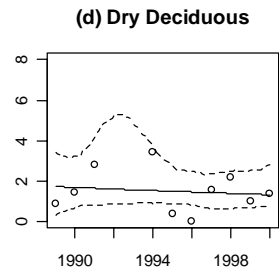
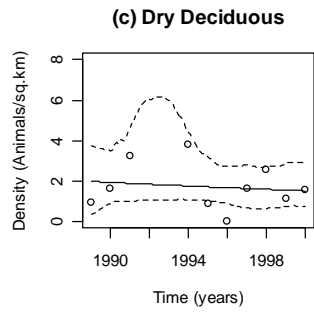
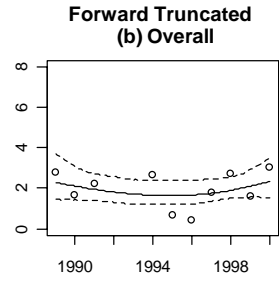
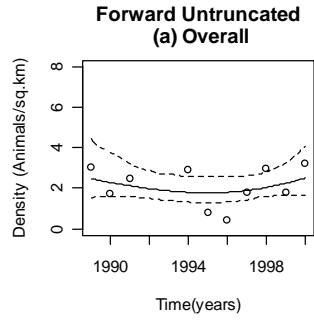
Appendix 4

Trends by habitat stratum for wild pig, elephant and gaur.

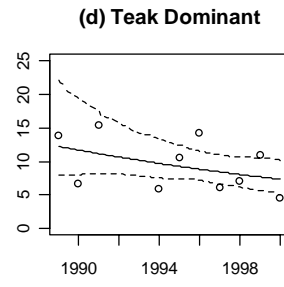
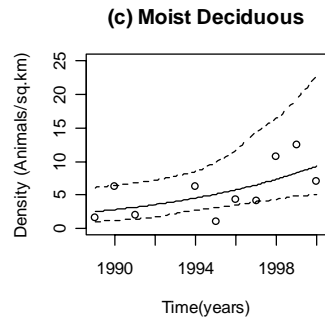
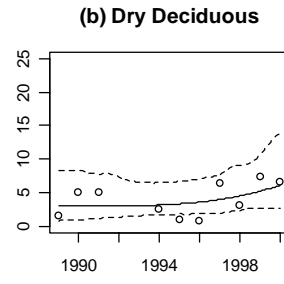
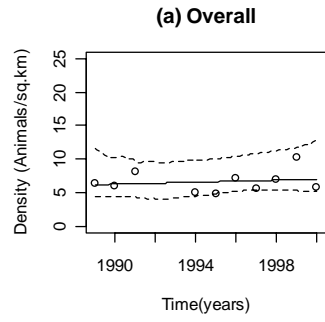
Wild Pig



Elephant



Gaur



Appendix 5

R code for bootstrap

```
bootstrap.data<-function(B,datafile="original.data.txt",file.base="boot",w,smooth){
#Created by:      Tiago Marques
#Updated by Aditya to store the results in a matrix and return the confidence intervals

#Defining the file name
data.file.name<-paste(file.base,".data.txt",sep="")

#read the data in
dados<-read.table(datafile,header=T,sep="\t",fill=T)
repl.DD<-repl.MD<-repl.TF<-repl.overall<-matrix(0,nrow=smooth,ncol=B)
Density.DD<-Density.MD<-Density.TF<-Density.overall<-matrix(0,nrow=10,ncol=B)
m<-0
errors<-NULL
for(r in 1:B)
  {
    create.data.file.bootstrap(dados,data.file.name)          #creates the bootstrap data
file for D5
    results<-cds(file.base=file.base,ext.files=T,bootstrap=T,w=w)          #runs D5 on the file
created, using the appropriate cmd file
    if(results$run.status>2)
      {
        #if there was an error
        m<-m+1          #add 1 to the number of errors
        errors[m]<-results$status
print("error")          #while keeping a track of what these errors are
      } else
      {
        write.table(results$gam.boot.edf,file="gam.boot.edf.xls",sep="\t",append=T)  #save these
as xls files
        write.table(results$gam.boot.overall.edf,file="gam.boot.overall.edf.xls",sep="\t",append=T)
        write.table(results$gam.boot.signif,file="gam.boot.signif.xls",sep="\t",append=T)

write.table(results$gam.boot.overall.signif,file="gam.boot.overall.signif.xls",sep="\t",append=T)
#for each resample
for(p in 1:smooth)
  {
    repl.DD[p,r]<-results$replicates.DD[p]
    repl.MD[p,r]<-results$replicates.MD[p]
    repl.TF[p,r]<-results$replicates.TF[p]
    repl.overall[p,r]<-results$replicates.overall[p]
  }

  for(q in 1:10)
  {
    Density.DD[q,r]<-results$D.DD[q]
    Density.MD[q,r]<-results$D.MD[q]
    Density.TF[q,r]<-results$D.TF[q]
    Density.overall[q,r]<-results$D.overall[q]
  }

  }
print(r)
}
lcl.repl.DD<-ucl.repl.DD<-lcl.repl.MD<-ucl.repl.MD<-lcl.repl.TF<-ucl.repl.TF<-lcl.repl.overall<-
ucl.repl.overall<-numeric(smooth)
```

```

for(p in 1:smooth)
{
  repl.sorted.DD<-repl.DD[p,]
  repl.sorted.DD<-sort(repl.sorted.DD)
  repl.sorted.MD<-repl.MD[p,]
  repl.sorted.MD<-sort(repl.sorted.MD)
  repl.sorted.TF<-repl.TF[p,]
  repl.sorted.TF<-sort(repl.sorted.TF)
  lcl.repl.DD[p]<-repl.sorted.DD[5]
  ucl.repl.DD[p]<-repl.sorted.DD[295]
  lcl.repl.MD[p]<-repl.sorted.MD[5]
  ucl.repl.MD[p]<-repl.sorted.MD[295]
  lcl.repl.TF[p]<-repl.sorted.TF[5]
  ucl.repl.TF[p]<-repl.sorted.TF[295]
  repl.overall.sorted<-repl.overall[p,]
  repl.overall.sorted<-sort(repl.overall.sorted)
  lcl.repl.overall[p]<-repl.overall.sorted[5]
  ucl.repl.overall[p]<-repl.overall.sorted[295]
}
lcl.Density.DD<-lcl.Density.MD<-lcl.Density.TF<-ucl.Density.DD<-ucl.Density.MD<-
ucl.Density.TF<-lcl.Density.overall<-ucl.Density.overall<-rep(0,10)
for(q in 1:10)
{
  Density.sorted.DD<-Density.DD[q,]
  Density.sorted.DD<-sort(Density.sorted.DD)
  Density.sorted.MD<-Density.MD[q,]
  Density.sorted.MD<-sort(Density.sorted.MD)
  Density.sorted.TF<-Density.TF[q,]
  Density.sorted.TF<-sort(Density.sorted.TF)
  lcl.Density.DD[q]<-Density.sorted.DD[5]
  ucl.Density.DD[q]<-Density.sorted.DD[295]
  lcl.Density.MD[q]<-Density.sorted.MD[5]
  ucl.Density.MD[q]<-Density.sorted.MD[295]
  lcl.Density.TF[q]<-Density.sorted.TF[5]
  ucl.Density.TF[q]<-Density.sorted.TF[295]
  Density.overall.sorted<-Density.overall[q,]
  Density.overall.sorted<-sort(Density.overall.sorted)
  lcl.Density.overall[q]<-Density.overall.sorted[5]
  ucl.Density.overall[q]<-Density.overall.sorted[295]
}
return(list(repl.DD=repl.DD,repl.MD=repl.MD,repl.TF=repl.TF,repl.overall=repl.overall,Density.DD=
Density.DD,Density.MD=
Density.MD,Density.TF=Density.TF,Density.overall=Density.overall,errors=errors,lcl.Density.DD=lcl.
Density.DD,lcl.Density.MD=lcl.Density.MD,lcl.Density.TF=lcl.Density.TF,ucl.Density.DD=ucl.Densi
ty.DD,ucl.Density.MD=ucl.Density.MD,ucl.Density.TF=ucl.Density.TF,
lcl.repl.DD=lcl.repl.DD,lcl.repl.MD=lcl.repl.MD,lcl.repl.TF=lcl.repl.TF,ucl.repl.DD=ucl.repl.DD,ucl.r
epl.MD=ucl.repl.MD,ucl.repl.TF=ucl.repl.TF,lcl.Density.overall=lcl.Density.overall,ucl.Density.overal
l=ucl.Density.overall,lcl.repl.overall=lcl.repl.overall,ucl.repl.overall=ucl.repl.overall))
}

```

#BOOTSTRAP PROCEDURE

```

create.data.file.bootstrap<-function(data,filename="boot.data.txt"){
#Created by:      Tiago Marques

```

```

#Updated by Aditya to resample by temporal replicates within each habitat stratum

```

```

#Purpose:      This function is called by the function bootstrap.data and creates a resample data
file,

```

```

#      based on an original file created by Distance

```

```

#Inputs:      data - a data.frame with the data as created by D5
#            filename - the name of the data file to be created

#Returns:     nothing, but creates a file in the current directory, named file.base+"data.txt", to be
              analysed by Distance
str.labels<-levels(as.factor(data$Year))           #an object with the strata labels
str.n<-length(str.labels)                         #counts the number of strata
#the strata ID numbers
sampler.labels<-NULL
temporal.labels<-NULL                            #an object to hold the sampler labels
file.create(filename)
file.line<-1                                     #a counter for the line number (it's the sampler ID
in the new file )
for(j in 1:str.n) {                               #for each strata
  sampler.labels<-levels(as.factor(data$Habitat.stratum[data$Year==str.labels[j]])) #get the
sampler labels
  n.sampler.labels<-length(sampler.labels)        #counts the samplers
in the strata
  for(k in 1:n.sampler.labels){                  #for each
sampler
  temporal.labels<-
levels(as.factor(data$Replicate[data$Habitat.stratum==sampler.labels[k]&data$Year==str.labels[j]]))
  #define temporal labels
  n.temporal.labels<-length(temporal.labels)      #counts number of
temporal replicates
  resample.temporal<-sample(x=temporal.labels,size=n.temporal.labels,replace=T) #resample the
temporal replicates within each line
  for(i in 1:n.temporal.labels) {                #for each resampled temporal replicate
    temp<-data[data$Year==str.labels[j]          &
data$Habitat.stratum==sampler.labels[k]&data$Replicate==resample.temporal[i,] #select it's
corresponding data
    temp.lines<-length(temp$Stratum.ID)          #count number of
lines in temp
    for(m in 1:temp.lines){                      #for each of these
lines
if(is.na(temp$PD[temp.lines])) {                #if no animals detected in this sampler, send
that info to data file
  cat(paste(1,".",sep=""),0,paste(temp[m,]$Stratum.ID,"
",str.labels[j],file.line,sep=""),temp[m,]$Effort,"\t","\t","\t", sep="\t", "\n", file=filename,
append=TRUE)
  file.line<-file.line+1 }                      #increase the sampler ID

  else {
    cat(paste(1,".",sep=""),0,paste(temp[m,]$Stratum.ID,"
",str.labels[j],sep=""),temp[m,]$Effort,temp[m,]$PD,temp[m,]$Cluster.size,as.character(temp[m,]$Hab
itat.Type),sep="\t", "\n", file=filename, append=TRUE)
    if(m==temp.lines) {                          #after the last animal of the sampler
      file.line<-file.line+1 } } } }            #increase the sampler ID
    } }
return()
}

```

```

cds <- function (key, adj, L, w, A=NA, xi, zi, file.base,ext.files=F,bootstrap=F) {
#Purpose: Driver function to run the mcdfs.exe engine from R
#Updated by Tiago on 19/12/2004 so that it can take an external data and command input files

```

```

#Inputs:
# key - vector string containing key functions

```

```

# adj - vector string containing adjustment terms
# L - total survey effort
# A - survey area
# xi - vector of perpendicular distances
# zi - vector of cluster sizes
# file.base - if specified and engine is cds, then the cds input and
# output files are written into the current directory, with the
# cds.file.base as a prefix and '.txt' as a suffix. E.g., setting
# cds.file.base to 'cds' produces 'cds.cmd.txt', 'cds.data.txt',
# 'cds.log.txt','cds.stat.txt','cds.plot.txt' and 'cds.boot.txt'
# If not specified, these files are created in a temp location and
# are deleted at the end.
# ext.files - If true, then the function uses the external files 'file.base'+'.cmd'+'.txt' as command file
# and 'file.base'+'.data'+'.txt' as data file. Typically these files would be the result of
# running Distance in debug mode, and should be placed in the working directory for R.
# By default ext.files is false, so the function looks for the files produced by functions
# 'create.data.file' and 'create.command.file'
# You need to change the input command file in order for the mcads engine to produce the files
# that the function 'read.stats.file' expects, and that means that inside the command file you
should define
# file names with prefix = file.base
# bootstrap - If true, procedure is being called inside a bootstrap routine, and intermediate files
# are deleted at each loop step, except the command file
#Outputs: list, containing
# densities by habitat stratum and overall densities, and degrees of freedom and significance
# Note - status is an integer:
# 1=OK, 2=warnings, 3=errors, 4=file errors, 5=some other problem (e.g., program crash)

#get input and output file names
file.names<-get.cds.file.names(file.base)

#if external not provided, create data file
if(ext.files==F) {
  create.data.file(file.names$data.file, L, A, xi, zi)
}

#if external not provided,create command file
if(ext.files==F) {
  create.command.file(file.names, key, adj, w)
}

#call cds engine
run.status<-run.cds(file.names[1])

#harvest results from stats file
params <- read.stats.file(file.names$stat.file)
res<-density.estimates(file="boot.data.txt",parameters=params$parameters,w=w)
#remove any temp files
if(bootstrap==F) {
  if(is.null(file.base)) remove.files(file.names)
}
else {#if inside a bootstrap routine, always delete intermediate files, except the command file
  remove.files(file.names,bootstrap=bootstrap)
}
#return results
return(list(run.status=run.status,years=res$years, habitat.strat=res$habitat.strat,
D=res$D,D.overall=res$D.overall,

replicates=res$replicates,replicates.overall=res$replicates.overall,replicates.DD=res$replicates.DD,replicates.MD=res$replicates.MD,replicates.TF=res$replicates.TF,

```



```
D.DD=res$D.DD,D.MD=res$D.MD,D.TF=res$D.TF,gam.boot.edf=res$gam.boot.edf,gam.boot.overall
.edf=res$gam.boot.overall.edf,gam.boot.signif=res$gam.boot.signif,gam.boot.overall.signif=res$gam.b
oot.overall.signif))
```

```
}
```

```
get.cds.file.names<-function(file.base=NULL) {
#Purpose: returns a list of filenames given the file.base
# If file.base is NULL filenames are in a temp directory
# Otherwise, they are of the form file.base +
# '.cmd'/'.data','.out','.log','.stat','.boot','.plot' + '.txt'

files.mid<-c("cmd","data","out", "log","stat","boot","plot")
files.suffix<-".txt"

if(is.null(file.base)) {
#get temp file names
file.names<-tempfile(files.mid)
} else {
file.names<-rep("",length(files.mid))
#and add the file base
for(i in 1:length(files.mid)) {
file.names[i]<-paste(file.base,".",files.mid[i],sep="")
}
}
#add the appropriate suffix
for(i in 1:length(files.mid)) {
file.names[i]<-paste(file.names[i],files.suffix,sep=".")
}

return(list(cmd.file=file.names[1], data.file=file.names[2], out.file=file.names[3],
log.file=file.names[4], stat.file=file.names[5], boot.file=file.names[6],
plot.file=file.names[7]))
}

run.cds<-function(cmd.file.name) {
#Purpose: runs the MCDS.exe engine and waits for it to finish
# *Note* that mcids.exe needs to be in the working directory, or in the
# PATH windows environment
# variable for this to work, as it makes no attempt to find the
# location of the file
#Inputs:
# cmd.file.name - name of the command file to run
#Returns:
# A status integer - 1=OK, 2=warnings, 3=errors, 4=file errors, 5=some other problem (e.g., program
crash)

command <- paste("mcids 0, \\"", cmd.file.name, "\"",sep="")
res<-system(command, intern=TRUE, invisible=TRUE)
#this line added by ljt 18/8/05 - FORTRAN sometimes returns lots of lines
if(is.vector(res)) res <- res[1]
res<-as.integer(res)
if(is.na(res)) res<-5
return(res)
}
```

```

read.stats.file<-function(stat.file.name) {
# Author: Ingrid

#changed by Aditya to give parameter estimates
# Purpose: extracts results statistics from the MCDS stat file

#Input:
# stat.file.name - name of file to look in

#Returns: list:
# parameters - parameter estimates
#read the file in and test that it has something in it
lines.v<-readLines(stat.file.name)
n.lines<-length(lines.v)
if(n.lines==0) {
  stop ("Nothing to read!")
}
#go through each line, parsing it and looking
# for the results we want
#and storing them in the appropriate place
parameters<-NULL
for (line in 1:n.lines)
  {
    parsed.line<-split.line(lines.v[line])
    if(parsed.line$ok)
      {
        if (parsed.line$module==2&parsed.line$statistic==101){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==102){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==103){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==104){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==105){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==106){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==107){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==108){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==109){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==110){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==111){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==112){
          parameters[length(parameters)+1]<-parsed.line$value }
        if (parsed.line$module==2&parsed.line$statistic==113){
          parameters[length(parameters)+1]<-parsed.line$value }
      }
  }
return(list(parameters=parameters))
}

density.estimates<-function(file="boot.data.txt",parameters,w){
#author:aditya

```

```

#to get density estimates and do a GAM on these
data<-
read.table(file,fill=T,sep=" ",header=F,na.strings=" ",blank.lines.skip=FALSE,col.names=c("V1","V2","
V3","V4","V5","V6","V7","V8","V9","V10","V11","V12"))
A.1<-parameters[1];A.2<-NULL;A.3<-NULL;A.4<-NULL;A.5<-NULL;A.6<-NULL;A.7<-
parameters[7];A.8<-parameters[8]
num.lines<-length(data$V2)          #set up a data frame for results
sigma<-rep(0,num.lines)
g<-rep(0,num.lines)
p<-rep(0,num.lines)
ht<-rep(0,num.lines)
Year<-rep(0,num.lines)
Habitat.stratum<-rep(0,num.lines)
Effort<-rep(0,num.lines)
solving<-
data.frame(sigma=sigma,g=g,p=p,ht=ht,Year=Year,Habitat.stratum=Habitat.stratum,Effort=Effort)

for(i in 1:num.lines)
{
  if(is.na(data[i,]$V6)|data[i,]$V6>w){i<-i+1}
  else{
    if(data[i,]$V8=="CL"){A.2<-parameters[2]}
    else {A.2<-0}

    if(data[i,]$V8=="DD"){A.3<-parameters[3]}
    else {A.3<-0}

    if(data[i,]$V8=="HD"){A.4<-parameters[4]}
    else{A.4<-0}

    if(data[i,]$V8=="MD"){A.5<-parameters[5]}
    else{A.5<-0}

    if(data[i,]$V8=="SF"){A.6<-parameters[6]}
    else{A.6<-0}
  }
#get s for each observation- the above will vary depending on the species.

  solving[i,1]<- A.1 * exp(A.2 + A.3 +A.4 +A.5 +A.6)

print(solving[i,1])
  integrand<-function(y,scale,adj1,adj2,w){(exp(-
y^2/(2*(scale^2)))*(1+((adj1*cos((2*pi*y)/w))+adj2*cos((3*pi*y)/w))))/(1+(adj1+adj2))}
  x<-integrate(integrand,lower=0,upper=w,scale=solving[i,1],adj1=A.7,adj2=A.8,w=w)
#get detection probaibility
  solving[i,2]<-x$value
  solving[i,3]<-solving[i,2]/w
  solving[i,4]<-data[i,]$V7/solving[i,3]
  solving[i,5]<-data[i,]$V4
  if(data[i,]$V3==1|data[i,]$V3==4|data[i,]$V3==7|data[i,]$V3==10|data[i,]$V3==13|data[i,]$
V3==16|data[i,]$V3==19|data[i,]$V3==22|data[i,]$V3==25|data[i,]$V3==28){solving[i,6]<-"DD"}
  if(data[i,]$V3==2|data[i,]$V3==5|data[i,]$V3==8|data[i,]$V3==11|data[i,]$V3==14|data[i,]$
V3==17|data[i,]$V3==20|data[i,]$V3==23|data[i,]$V3==26|data[i,]$V3==29){solving[i,6]<-"MD"}
  if(data[i,]$V3==3|data[i,]$V3==6|data[i,]$V3==9|data[i,]$V3==12|data[i,]$V3==15|data[i,]$
V3==18|data[i,]$V3==21|data[i,]$V3==24|data[i,]$V3==27|data[i,]$V3==30){solving[i,6]<-"TF"}
  #fill in data frame
  solving[i,7]<-data[i,]$V5
}
}
year<-c(1,2,3,6,7,8,9,10,11,12)
n.year<-length(year)

```

```

habitat<-levels(as.factor(c("DD","MD","TF")))
habitat.s<-c("DD","MD","TF")
n.habitat<-length(habitat)
D<-years<-habitat.strat<-rep(0,30)
for(k in 1:n.year){
  for(h in 1:n.habitat){
    m<-((k-1)*n.habitat+h)
    habitat.strat[m]<-habitat.s[h]
    years[m]<-year[k]+1988
    temp<-sum(solving$ht[solving$Year==year[k]
solving$Habitat.stratum==habitat.s[h]]/(solving$Effort[solving$Year==year[k]
solving$Habitat.stratum==habitat[h]]*2*(w/1000))
    if(length(temp)==0) {
      D[m]<-0
    } else {
      D[m]<-temp
    }
  }
}
}

```

```

A<-c(rep(c(30,30,40),10))
DA<-D*A
estimates<-data.frame(years=years,habitat.strat=habitat.strat,D=D,A=A,DA=DA)

```

```

D.DD<-estimates$D[estimates$habitat.strat=="DD"]
D.MD<-estimates$D[estimates$habitat.strat=="MD"]
D.TF<-estimates$D[estimates$habitat.strat=="TF"]
D.overall<-
c(sum(estimates$DA[estimates$years==1989])/100,sum(estimates$DA[estimates$years==1990])/100,s
um(estimates$DA[estimates$years==1991])/100,sum(estimates$DA[estimates$years==1994])/100,su
m(estimates$DA[estimates$years==1995])/100,
sum(estimates$DA[estimates$years==1996])/100,sum(estimates$DA[estimates$years==1997])/100,su
m(estimates$DA[estimates$years==1998])/100,sum(estimates$DA[estimates$years==1999])/100,sum(
estimates$DA[estimates$years==2000])/100)

```

#select densities for each habitat and for overall site

```

fac.1<-as.numeric(estimates$habitat.strat=="DD")
fac.2<-as.numeric(estimates$habitat.strat=="MD")
fac.3<-as.numeric(estimates$habitat.strat=="TF")

```

#set up dummy variables

```

gam.boot<-
gam(D~habitat.strat+s(years,by=fac.1,k=4)+s(years,by=fac.2,k=4)+s(years,by=fac.3,k=4),family=quasi
poisson,data=estimates)

```

do a GAM on the densities by habitat stratum

```

summary.gam.boot<-summary(gam.boot)
gam.boot.edf<-summary.gam.boot[15]
gam.boot.signif<-summary.gam.boot[8]
pred.year<-
c(rep(1989,3),rep(1989.1,3),rep(1989.2,3),rep(1989.3,3),rep(1989.4,3),rep(1989.5,3),rep(1989.6,3),rep(
1989.7,3),rep(1989.8,3),rep(1989.9,3),rep(1990,3),rep(1990.1,3),rep(1990.2,3),rep(1990.3,3),rep(1990.
4,3),rep(1990.5,3),rep(1990.6,3),rep(1990.7,3),rep(1990.8,3),rep(1990.9,3),
rep(1991,3),rep(1991.1,3),rep(1991.2,3),rep(1991.3,3),rep(1991.4,3),rep(1991.5,3),rep(1991.6,3),rep(1
991.7,3),rep(1991.8,3),rep(1991.9,3),rep(1992,3),rep(1992.1,3),rep(1992.2,3),rep(1992.3,3),rep(1992.4
,3),rep(1992.5,3),rep(1992.6,3),rep(1992.7,3),rep(1992.8,3),rep(1992.9,3),rep(1993,3),rep(1993.1,3),re

```

```
p(1993.2,3),rep(1993.3,3),rep(1993.4,3),rep(1993.5,3),rep(1993.6,3),rep(1993.7,3),rep(1993.8,3),rep(1993.9,3),rep(1994,3),rep(1994.1,3),rep(1994.2,3),
rep(1994.3,3),rep(1994.4,3),rep(1994.5,3),rep(1994.6,3),rep(1994.7,3),rep(1994.8,3),rep(1994.9,3),rep(1995,3),rep(1995.1,3),rep(1995.2,3),rep(1995.3,3),rep(1995.4,3),rep(1995.5,3),
rep(1995.6,3),rep(1995.7,3),rep(1995.8,3),rep(1995.9,3),rep(1996,3),rep(1996.1,3),rep(1996.2,3),rep(1996.3,3),rep(1996.4,3),rep(1996.5,3),rep(1996.6,3),rep(1996.7,3),rep(1996.8,3),rep(1996.9,3),
rep(1997,3),rep(1997.1,3),rep(1997.2,3),rep(1997.3,3),rep(1997.4,3),rep(1997.5,3),rep(1997.6,3),rep(1997.7,3),rep(1997.8,3),rep(1997.9,3),rep(1998,3),rep(1998.1,3),rep(1998.2,3),rep(1998.3,3),rep(1998.4,3),
rep(1998.5,3),rep(1998.6,3),rep(1998.7,3),rep(1998.8,3),rep(1998.9,3),rep(1999,3),rep(1999.1,3),rep(1999.2,3),rep(1999.3,3),rep(1999.4,3),rep(1999.5,3),rep(1999.6,3),rep(1999.7,3),rep(1999.8,3),
rep(1999.9,3),rep(2000,3),rep(2004,3))
```

```
#predict at these points
```

```
pred.habitat<-rep(c("DD","MD","TF"),112)
factor.1<-as.numeric(pred.habitat=="DD")
factor.2<-as.numeric(pred.habitat=="MD")
factor.3<-as.numeric(pred.habitat=="TF")
pred.frame<-
data.frame(years=pred.year,habitat.strat=pred.habitat,fac.1=factor.1,fac.2=factor.2,fac.3=factor.3)
```

```
replicates<-predict.gam(gam.boot,pred.frame,type="response")
```

```
pred.frame<-cbind(pred.frame,replicates=replicates)
```

```
replicates.DD<-pred.frame$replicates[pred.frame$habitat.strat=="DD"]
replicates.MD<-pred.frame$replicates[pred.frame$habitat.strat=="MD"]
replicates.TF<-pred.frame$replicates[pred.frame$habitat.strat=="TF"]
```

```
years.overall<-c(1989,1990,1991,1994,1995,1996,1997,1998,1999,2000)
estimates.overall<-data.frame(D.overall=D.overall,years.overall=years.overall)
```

```
gam.boot.overall<-gam(D.overall~s(years.overall,k=4),family=quasipoisson,data=estimates.overall)
```

```
#set up gam for the overall density
```

```
summary.gam.boot.overall<-summary(gam.boot.overall)
```

```
gam.boot.overall.edf<-summary.gam.boot.overall[15]
gam.boot.overall.signif<-summary.gam.boot.overall[8]
pred.year.overall<-c(seq(1989,2000,.1),2004)
pred.frame.overall<-data.frame(years.overall=pred.year.overall)
```

```
replicates.overall<-predict.gam(gam.boot.overall,pred.frame.overall,type="response")
```

```
return(list(D.overall=D.overall,D=D,D.DD=D.DD,D.MD=D.MD,D.TF=D.TF,replicates=replicates,replicates.DD=replicates.DD,replicates.MD=replicates.MD,replicates.TF=replicates.TF,replicates.overall=replicates.overall,gam.boot.edf=gam.boot.edf,gam.boot.overall.edf=gam.boot.overall.edf,gam.boot.signif=gam.boot.signif,gam.boot.overall.signif=gam.boot.overall.signif))
}
```

```
split.line <- function(line) {
#takes each line and returns the different values for stratum, samp, estimator, ...
if(nchar(line)<6) stop ("This isn't a stats file!")
stratum<-as.integer(substr(line,2,6))
if(is.na(stratum)) {
#this means that
ok=FALSE
}
```

```

samp<-NULL
estimator<-NULL
module<-NULL
statistic<-NULL
value<-NULL
cv<-NULL
lcl<-NULL
ucl<-NULL
degrees.freedom<-NULL
} else {
ok=TRUE
#parse (understand!) the rest of the line when ok is true
samp <- as.integer(substr(line, 8, 12))
estimator<-as.integer(substr(line,14,14))
module<-as.integer(substr(line, 16,16))
statistic<-as.integer(substr(line,18,20))
value<-as.numeric(substr(line, 22,35))
cv<-as.numeric(substr(line, 37,50))
lcl<-as.numeric(substr(line, 52,65))
ucl<-as.numeric(substr(line, 67 ,80))
degrees.freedom<-as.integer(substr(line, 82,95))
}
return(list(ok=ok,stratum=stratum,samp=samp,
estimator=estimator, module=module, statistic=statistic,
value=value, cv=cv, lcl=lcl, ucl=ucl, degrees.freedom=degrees.freedom))
}

```

```

remove.files<-function(file.names,bootstrap=F) {
#Purpose: removes the files listed in the vector file.names
#Written by: Len Thomas 26/7/04
#Updated by: Tiago Marques 01-01-2005 so that the command file is not deleted (if used inside
bootstrap loop)
start<-1
if(bootstrap==T) {
start<-2#if used inside bootstrap loop, the command file is not deleted
}
for (i in start:length(file.names)) {
if (file.exists(file.names[[i]])) file.remove(file.names[[i]])
}
}
}

```