Geo-information Science and Remote Sensing

Thesis Report GIRS-2026-12

INCORPORATING VIEWSHEDS INTO WILDLIFE ABUNDANCE ESTIMATION

Thijs Vons

March 17, 2026

WAGENINGEN
UNIVERSITY & RESEARCH

Incorporating Viewsheds into Wildlife Abundance Estimation

Thijs Vons

Registration number: 1583379

Supervisors:

dr. JAJ (Jasper) Eikelboom
ir. AR (Aldo) Bergsma
dr. Maya Beukes (External, Senckenberg)

A thesis submitted in partial fulfillment of the degree of Master of Science
at Wageningen University and Research, The Netherlands.

March 17, 2026
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2026-12
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Acknowledgements

I would like to thank my supervisors, Jasper Eikelboom and Aldo Bergsma, for their guidance throughout this thesis. Through their feedback, brainstorming sessions, and open discussions, they helped me shaping this research. Although I was not very familiar with the field of ecology prior to this thesis, they supported me in applying my skills to this fascinating domain.

I would also like to thank Maya Beukes for initiating this project and for always being open to discussion. I am grateful to her and her team for taking the time during their own research in South Africa to conduct game counts that contributed to the data used in this thesis.

My thanks also go to GeoSMART for providing the 5 by 5 metre Digital Elevation Model of the study area. Without this dataset it would not have been possible to apply my method to the obtained game counts.

Finally, I would like to thank the other students in the thesis room for the many fruitful discussions, as well as the even more productive breaks filled with table tennis.

Abstract

Due to anthropogenic activities, such as climate change and overexploitation of species and land, the global rate of species extinction has been accelerating, leading to a decrease in wildlife abundance. To investigate the effect of anthropogenic activities on wildlife abundance, it is crucial to have insight into the trends in abundance of various species. A commonly applied method for estimating wildlife abundance is distance sampling. Rather than directly using animal counts for abundance estimation, distance sampling uses the distance from an observer to the animal to provide a more accurate abundance estimate by estimating the detection probability of detected animals based on the distance distribution. However, distance sampling does not take into account the heterogeneity of the terrain.

This study aimed to incorporate terrain heterogeneity into distance sampling by incorporating viewsheds. I developed viewshed compensated distance sampling, a method that explicitly uses the viewshed to compensate for the distance distribution of detected animals for a more accurate abundance estimate. My method was based on simulated viewsheds and simulated animal populations, as this enables validation of the method across many animal populations. The performance of my method was compared to the performance of traditional distance sampling. Additionally, my method was compared to a previously proposed method incorporating viewsheds into distance sampling using Bayesian statistics, by modifying the probability density function of the sampled area. Finally, I applied my method to the real viewshed of the Nossob riverbed and a real animal population in the study area of the Kgalagadi Transfrontier Park in South Africa. With the real viewshed of the Nossob riverbed and simulated animal populations with a true abundance of 1.000, the abundance estimate improves from 502 ± 37 animals for traditional distance sampling to 990 ± 158 animals for my method. With the real viewshed of the Nossob riverbed and the real animal population the abundance estimate by my method is $167.0\% \pm 42.7\%$ larger than the abundance estimate by traditional distance sampling. My method improves on the Bayesian distance sampling method by estimating abundance for the full area of interest, rather than estimating abundance for the sampled area only. My method allows better insight into trends in wildlife abundance by more accurately estimating wildlife abundance.

Contents

Glossary

Table 1: Explanation of important terms and abbreviations used in this thesis.

Term	**Definition**
Baseline distance sampling	Viewshed incorporated distance sampling method fitting the distance distribution of detected animals before it is viewshed compensated.
Bayesian distance sampling	Viewshed incorporated distance sampling method using Bayesian statistics.
DEM	Digital Elevation Model.
Distance sampling	Abundance estimation method using the distance from an observer to the animal to provide a more accurate abundance estimate by estimating the detection probability of detected animals based on the distance distribution.
Fitting viewshed parameter distance sampling	Viewshed incorporated distance sampling method where the detection function is modified to include a viewshed parameter.
KTP	Kgalagadi Transfrontier Park. Located in South Africa. Study area of this thesis.
Game counts	Systematic method to determine wildlife abundance in a predefined area.
Mixed input data	Input data where the viewshed is real and the animal population has been simulated.
Nossob viewshed	Viewshed of the Nossob riverbed.
PDF	Probability density function.
Real input data	Input data where both the viewshed and the animal population are real.
Simulated input data	Input data where both the viewshed and the animal population have been simulated.
Traditional distance sampling	Distance sampling method fitting the distance distribution of detected animals
Viewshed	The geographical area that is visible from a specific point.
Viewshed compensated distance sampling	Viewshed incorporated distance sampling method fitting the distance distribution of detected animals after it is viewshed compensated.
Viewshed incorporated distance sampling	Any distance sampling method that incorporates viewsheds.

1 Introduction

The global rate of species extinction is already at least tens to hundreds of times higher over the past centuries than the average rate over the past 10 million years and the rate is accelerating (IPBES et al., 2019). This has led to a decrease in wildlife abundance. Wildlife abundance is defined as the total sum of individuals from a given species within a given area (Dubey, 2023). From 1970 to 2005, wildlife abundance has been decreasing, even in protected areas (Craigie et al., 2010). The main reason for the decline in wildlife abundance is anthropogenic activities (Okorondu et al., 2022) such as climate change, overexploitation of species and land, invasive species and diseases (Prakash and Verma, 2022).

To investigate the effect of anthropogenic activities on wildlife abundance, it is crucial to have an insight into the trends in the abundance of various species. Since it is not feasible to count every individual animal, abundance estimation methods have been widely applied. According to Callaghan et al. (2024), common quantitative methods of abundance estimation include capture-mark-recapture (White, 1982), N-mixture modelling (Royle, 2004), relative abundance indices and distance sampling (Buckland et al., 2004). Other methods available are removal sampling, point transect sampling, line transect sampling and aerial count (Norton-Griffiths, 1978). These methods often involve repeated measurements of the same study area, and direct contact with wildlife, to mark, track or remove them. Ideally, repeated measurements and direct contact is minimised.

All of these methods are based on assumptions, which are important to satisfy. Transect sampling methods are based on 6 assumptions (the last two only apply to distance sampling) (Batter et al., 2022; Buckland et al., 2004; Gates et al., 1968):

1. Animals on the transect (i.e. distance = 0) are detected with certainty.
2. All animals are counted accurately.
3. Animals are distributed independently of the transect.
4. All animals within the sampling area must be available for detection.
5. Distance and bearing are measured to the location where animals naturally occur.
6. Distance and bearing are measured accurately.

Assumptions 1, 2, 5 and 6 are within control of the observer, whereas assumptions 3 and 4 cannot directly be controlled in the field. To avoid violation of these assumptions, fieldwork protocols should aim to satisfy these assumptions. However, it is not feasible to fully satisfy all assumptions this way. When the terrain is heterogeneous, with changing vegetation, diverse elevation profiles and different land cover, assumptions 3 and 4 are violated. The lack of visibility of animals impacts the performance of abundance estimation methods. Therefore, it is necessary to modify the methods to ensure accurate abundance estimation. Proposed modifications include the implementation of covariate models (Buckland et al., 2004, 2015), Horvitz-Thompson-like estimators/probability detection functions (Kansanen et al., 2021), or stratification (Buckland et al., 2015; Morgan et al., 2024).

Another proposed solution to relaxing assumptions related to heterogeneous terrain is to incorporate viewsheds. A viewshed predicts the total area visible from a single point (Fisher, 1991). A complete viewshed means there are no obstacles and an incomplete viewshed means that parts of the area are blocked by objects such that an observer can not see that area and/or the area behind the

object. Rather than directly implementing viewsheds into the transect sampling methods, Maichak and Schuler (2004) applied viewsheds to wildlife abundance estimation by investigating the viewshed of the area of interest and applying this knowledge in the planning of the line transect route. Driving routes with the most visible area gave significantly more accurate animal abundance estimation numbers compared to worse viewsheds or random point samples.

Research has been done on directly incorporating viewsheds into transect sampling methods. Koeck et al. (2025) applied a Space-to-Event (STE) model which uses time-lapse photography of unmarked animals to relate animal detection to viewsheds (developed by a camera sampling area). STE can be used to determine animal density estimates which can be extrapolated to an animal abundance estimation over large areas. Similarly, Carswell et al. (2025) has applied Realised Viewshed Size (RVS) in camera-trap-based abundance estimation methods to capture how probability changes due to heterogeneous terrain and other influences. These are both examples of point sampling methods using cameras.

Not much research has been done on directly incorporating viewsheds into transect sampling methods. However, in the domain of distance sampling, Delaney et al. (2025) published a paper on the topic of incorporating viewsheds into distance sampling. This paper defines the probability density function (PDF) of the sampling area, to incorporate viewsheds into the method. The PDF is a critical part of the distance sampling method, describing the estimated probability density of encountering an animal at a certain distance. Delaney et al. (2025) allowed for a modification of the PDF, by applying a novel Bayesian statistics distance sampling method rather than applying the commonly used Distance Sampling R package (Thomas et al., 2010). Delaney et al. (2025) simulated three datasets with different viewsheds to test the method. Both the traditional and customised models were applied to compare parameter estimates using a dataset of 95 county-level spotlight surveys of white-tailed deer (Odocoileus virginianus) in a study area in Iowa, USA (Kaminski et al., 2019). The results show that customising the PDF improves animal density estimates and offers flexibility when it cannot be assumed that animals are randomly distributed.

Delaney et al. (2025) introduced a basic method for incorporating viewsheds into distance sampling, but it has several limitations. The first two apply more broadly to transect sampling methods. (1) Only a single PDF is constructed for the entire area of interest. This approach does not account for the sight-line from the observer to the animal at the time of observation, which is especially important to viewsheds. (2) An animal which has been detected may not be visible from a perpendicular line to the transect, while the distance used for distance sampling is taken along this line. (3) Abundance is estimated for the sampled area only. To obtain the estimate for the full area, the abundance estimate has to be extrapolated from the sampled area to the full area.

While the research above indicates substantial progress in relaxing assumptions 3 and 4 to account for heterogeneous terrain, there is still limited research on direct incorporation of viewsheds into transect sampling models, particularly given the specific limitations outlined above.

The aim of this research was to develop a method that incorporates viewsheds into distance sampling (viewshed incorporated distance sampling) and improves on the previously proposed method by Delaney et al. (2025).

The performance of my method was evaluated according to the following research objectives, concerned with pairwise comparison of the performance between my method, traditional distance sampling and the method proposed by Delaney et al. (2025), across different viewshed and animal population input data combinations:

Research objective 1:
Comparing my method with a simulated viewshed and a simulated animal population (simulated input data) to:

- Traditional distance sampling.
- The method by Delaney et al. (2025).

Research objective 2:
Comparing my method with a real viewshed and a simulated animal population (mixed input data) to:

- Traditional distance sampling.
- My method with simulated viewsheds and a simulated animal population.

Research objective 3:
Comparing my method with a real viewshed and a real animal population (real input data) to:

- Traditional distance sampling.
- My method with a real viewshed and a simulated animal population.

2 Methods

In this section first traditional distance sampling will be explained. Next, 3 methods incorporating viewsheds are explained, with one of these methods included in the appendix. (1) Firstly, the method proposed by Delaney et al. (2025), which for this thesis is called 'Bayesian distance sampling'. (2) Secondly, I modified Bayesian distance sampling. This resulted in multiple methods, which for this thesis are collectively called 'Modified Bayesian distance sampling'. (3) Thirdly, my method called 'Viewshed compensated distance sampling' for this thesis is explained. This is the main method of this thesis. After outlining these methods, research objective specific methods are explained. Another method I designed called 'Fitting viewshed parameter distance sampling', is explained in appendix A.3.

2.1 Distance sampling

Distance sampling is a method to approximate animal abundance where an observer travels along a line or to points, counting animals or animal clusters, while measuring the distance and bearing to animals. From the distance and bearing, the perpendicular distance of animals or animal clusters to the transect is calculated. Using the distance of animals or animal clusters an abundance estimate is made. For the rest of this thesis traditional distance sampling refers to the distance sampling described in this section. The following section is based on (Buckland et al., 2004).

First consider the non-distance sampling case where an observer drives a straight transect with length $L[m]$. It is assumed that all animals up to half-width $w[m]$ (width from transect to edge) are detected. In this case, the estimated animal abundance is $\hat{N} = n$. Where N is the abundance and n is the number of counted animals. The $\hat{}$ indicates that a quantity is an estimate based on sampled data.

Issues arise with this very basic method due to the assumption that all animals on the plot are detected:

- The more time spent on the plot by the observer, the greater the bias which arises from duplicate counts of animals entering or leaving the plot.
- For animals close to the boundary of the plot it can be hard to judge whether they are contained within the plot or not.
- It can be difficult to judge whether an observer has detected all animals in the plot.

Distance sampling tackles these issues by introducing two probability density functions. (1) The PDF of distances of detected and undetected animals, $\pi(x)$. Where $x\ [m]$ is an animals distance to the transect, with $0 \leq x \leq w$. For a straight transect this PDF is a uniform distribution, such that $\pi(x) = 1$. (2) The PDF of distances of detected animals, $f(x)$.

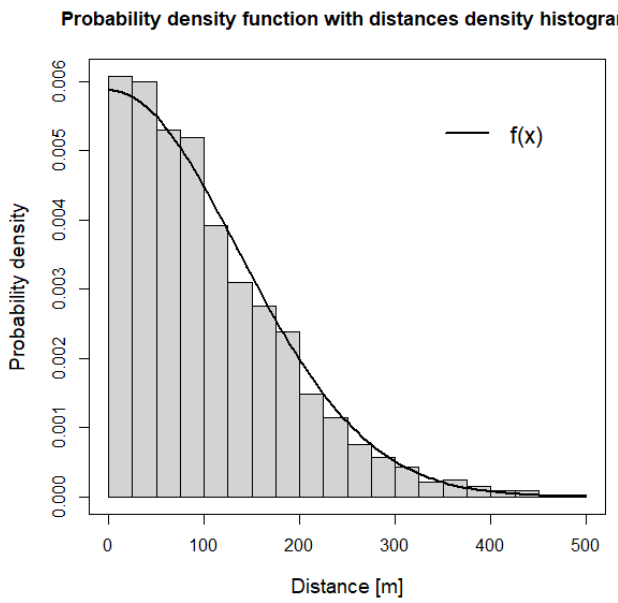
To model the PDF of distances of detected animals, the detection function $g(x)$ is defined. $f(x)$ is calculated by normalising the area under $g(x)$ to 1, as seen in Figure 2.1a:

$$f(x) = \frac{g(x)}{\int_0^w g(x)\,dx} \tag{2.1}$$

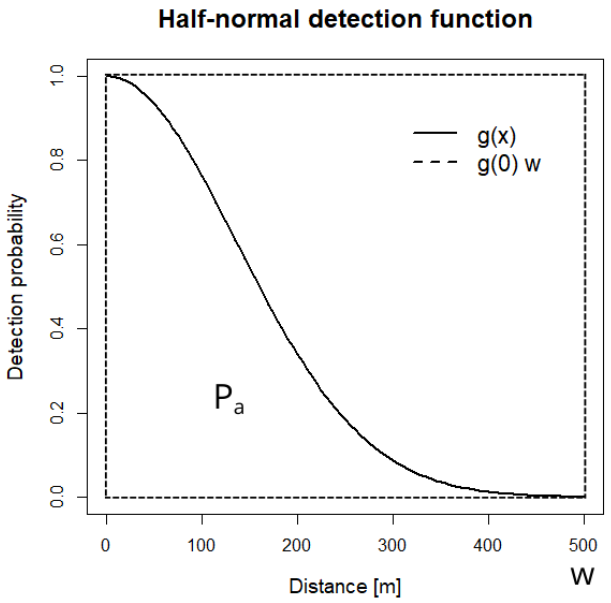
The detection probability, P_a, which is the proportion of animals detected, is visualised in Figure 2.1b and defined as:

$$\hat{P}_a = \frac{1}{w} \int_0^w g(x)\, dx \tag{2.2}$$

When it is assumed that all animals are detected, $g(x) = 1$, so $P_a = 1$. When this assumption does not hold, $0 \leq P_a \leq 1$.



(a) Example of probability density function $f(x)$, fitted through the probability density histogram of distances of detected animals. Fitted using maximum likelihood fitting. Area under the curve is 1.



(b) The detection function $g(x)$ is used to calculate probability detection P_a. P_a can be represented by the proportion of the dashed rectangle that is under the curve in this plot. Equation 2.2 applies. The area of the dashed rectangle is $g(0)\, w = w$, as $g(0) = 1$.

Figure 2.1: Visualisation of the probability density function $f(x)$ and the detection function $g(x)$.

The abundance estimate is calculated as:

$$\hat{N} = \frac{n}{\hat{P}_a} \tag{2.3}$$

A model for the detection function is required to calculate the detection probability. Since it is assumed that all animals on the transect are detected (assumption 1), the detection function has the restriction $g(0) = 1$. The half-normal equation is commonly used in distance sampling (Buckland et al., 2015), and therefore proposed as a model for $g(x)$:

$$g(x) = e^{-\frac{x^2}{2\sigma^2}} \tag{2.4}$$

σ is the scale parameter of the half-normal function and determines how quickly detection probability decreases with distance. The optimal σ based on the sampled distance data can be estimated using maximum likelihood fitting. The likelihood equation quantifies the probability of sampled data as a function of σ:

$$\mathcal{L}_x(\sigma) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{g(x_i)}{\int_0^w g(x)\, dy} \tag{2.5}$$

Taking the log:

$$\begin{aligned}\log \mathcal{L}_x(\sigma) &= \sum_{i=1}^{n} \left[\log g(x_i) - \log \int_0^w g(x)\, dx \right] \\ &= \sum_{i=1}^{n} \log g(x_i) - n \log \int_0^w g(x)\, dx\end{aligned} \tag{2.6}$$

This equation is minimised to find the optimal estimated σ, $\hat{\sigma}$. An example of a resulting fit with $\hat{\sigma}$ is shown in Figure 2.1a.

Combining $\hat{\sigma}$ and equations 2.2, 2.3 and 2.4:

$$\hat{N} = nw \left(\int_0^w e^{-\frac{x^2}{2\hat{\sigma}^2}}\, dx \right)^{-1} \tag{2.7}$$

This method generalises to areas larger than the sampled area, by calculating the animal density in the sampled area and extrapolating this to the full area of interest.

2.2 Bayesian distance sampling

The method proposed by Delaney et al. (2025) uses a Bayesian hierarchical model implemented in the R program JAGS (Plummer, 2003), with the R interface jagsUI (Kellner and Meredith, 2026).

A probability density function is defined that describes the proportion of the sampled area located at different distances from the observer. This is done by calculating the mean sampled area within each distance bin. It is assumed that animals only occur within the sampled area. This PDF is mathematically equivalent to $\pi(x)$ used in traditional distance sampling. For traditional distance sampling, the PDF is uniform. However, for an incomplete viewshed, the PDF is not uniform.

The aim of the model is to estimate two unknown parameters, σ and ψ, which best explain the distribution of distances of detected animals. σ is the scale parameter in the half-normal detection function (equation 2.4) and ψ represents the probability that an individual in an augmented dataset corresponds to a real animal in the population.

To estimate abundance, the model applies data augmentation. An augmented dataset of animals is created by combining the number of detected animals with a number of additional potential (but undetected) animals (M animals in total). For each animal in the augmented dataset, the model estimates a latent binary variable z_i that indicates whether the individual truly exists in the population. z_i follows a Bernoulli distribution with ψ as the probability of success (Dodge, 2008). A value of $z_i = 1$ indicates that the individual exists, whereas a value of $z_i = 0$ indicates that the individual does not exist.

σ and ψ are estimated using Markov Chain Monte Carlo (MCMC) sampling (Robert and Casella, 2011). The estimation begins with non-informative (flat) prior distributions: $\sigma \sim \mathrm{Uniform}(0, \sigma_{\max})$ and $\psi \sim \mathrm{Uniform}(0, 1)$. During the MCMC process, the algorithm repeatedly proposes values for σ and ψ and evaluates how well these values explain the ground truth data of detected animals, given the constraint of the PDF. Over many iterations, the initially broad prior distributions are updated to produce posterior distributions for the parameters. The resulting estimates are denoted $\hat{\sigma}$ and $\hat{\psi}$.

$\hat{\psi}$ determines how many individuals in the augmented dataset are estimated to be real animals, by defining z. The estimated population size is calculated as:

$$\hat{N} = \sum_{i=1}^{M} z_i \tag{2.8}$$

The abundance estimate $\hat{N}$ is not the abundance estimate for the full rectangular area, but for the sampled area only. The abundance estimate for the full area is obtained by extrapolating the abundance estimate of the sampled area, by dividing by the ratio between the sampled area and the full area.

2.3 Modified Bayesian distance sampling

The Bayesian distance sampling method proposed by (Delaney et al., 2025), only uses the absolute value of distances. Therefore, the distance bins of the PDF could only be defined from 0 to w. I modified the PDF such that the distance bins of the PDF extend to both sides of the transect, by defining distance bins for distances from $-w$ to w, such that $-w \leq x \leq w$. Since this modified PDF covers both negative and positive distance bins, the width of the distance bins is impacted. The width of the distance bins can be kept the same, such that double the number of distance bins is used. Alternatively, the width of the distance bins can be doubled, such that the number of distance bins is equal to the number of distance bins used in Bayesian distance sampling by Delaney et al. (2025).

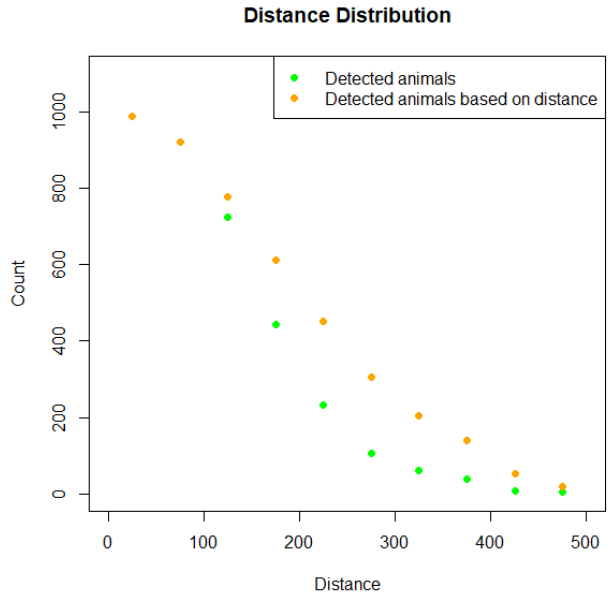
Taking it a step further, instead of averaging over an entire distance bin to calculate the PDF, a location specific PDF is defined for each animal. This is done by taking the sampled area in each distance bin at the x-coordinate of the animal. As with the previous modification, the width of the distance bins should be taken into account.

2.4 Viewshed compensated distance sampling

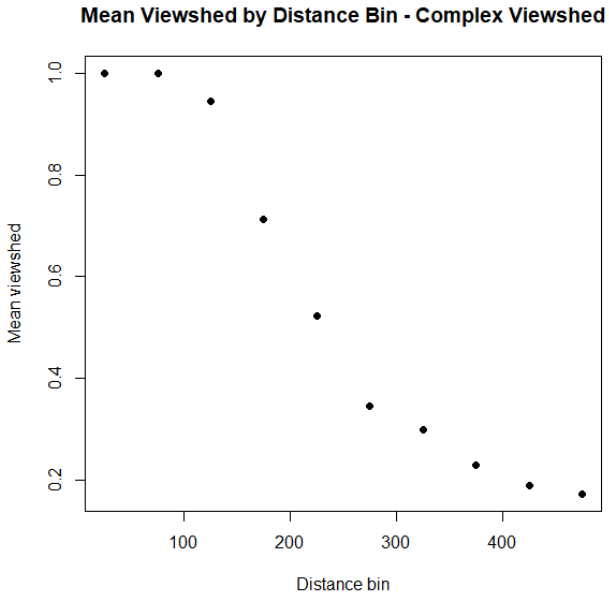
The following method, called viewshed compensated distance sampling, uses the distribution of viewshed parameters to compensate the distance distribution of detected animals. The viewshed parameter is an animal specific value, describing the likelihood that an observer was able to see the animal from the transect, with $v \in [0, 1]$.

For each detected animal the distance to the transect is extracted. These distances are counted per distance bin (example shown in Figure 2.2a in green).

To compensate the distance distribution of the detected animals for the effect of viewsheds, the mean viewshed parameter for each distance bin is calculated (example shown in Figure 2.2b). The detected animals count of each distance bin is divided by the mean viewshed parameter for that distance bin (dividing the green points by the black points). This operation removes the effect of the viewshed and results in the distribution if animal detection only depends on distance and not on the viewshed. This theoretical distribution (which is known in simulation) is shown in orange in Figure 2.2a. The aim of this method was to recreate this distribution. The compensated distance distribution alters the detection probability (P_a) and the number of counted animals (n). For this method n is defined as the sum of the values in the distance distribution.



(a) Distance count distributions. In green, the distribution of detected animals. In orange, the distribution of animals which were not detected, but would have been detected in a complete viewshed. Total population size = 10.000.



(b) Mean viewshed parameter of the complex viewshed for every distance bin.

Figure 2.2: Overview of viewshed compensated distance sampling method

As the next step, traditional distance sampling can be applied. However, traditional distance sampling takes as input a list of distances and finds the optimal σ value by minimising the negative log likelihood (section 2.1). This is not possible, since rather than a list of distances, a distribution of detected animal counts is available. However, this distribution allows for fitting of the detection function. For this thesis non-linear least squares (Teunissen, 1990) fitting to the half-normal equation (equation 2.4) was applied, with the input data being the green distribution in Figure 2.2a.

As absolute counts are used, $g(0) = 1$ does not hold. Therefore, the parameter τ, controlling the value of $g(x)$ at $x = 0$, is added to the original detection function:

$$g(x) = \tau e^{-\frac{x^2}{2\sigma^2}} \tag{2.9}$$

After fitting the data to get estimates for σ and τ, the detection probability is calculated. Due to the addition of τ to the detection function, equation 2.2 is modified with a factor $\frac{1}{\tau}$ resulting in:

$$\hat{P}_a = \frac{1}{\tau w} \int_0^w g(x)\, dx = \frac{1}{\tau w} \int_0^w \tau e^{-\frac{x^2}{2\sigma^2}} dx \tag{2.10}$$

τ can be pulled out of the integral and cancels out in the numerator and denominator. This leads to the final equation for the detection probability:

$$\hat{P}_a = \frac{1}{w} \int_0^w e^{-\frac{x^2}{2\hat{\sigma}^2}}\, dx \tag{2.11}$$

Equation 2.11 is equal to the original equation for detection probability for traditional distance sampling, equation 2.2.

The new number of counted animals is denoted as $\hat{n}$. Using $\hat{P}_a$ and $\hat{n}$, equation 2.7 is modified to calculate the abundance estimate.

$$\hat{N} = \hat{n}w \left(\int_0^w e^{-\frac{x^2}{2\hat{\sigma}^2}} \, dx \right)^{-1} \tag{2.12}$$

2.4.1 Incorporating cluster sizes

Cluster sizes were incorporated into this method in the step of counting animal distances per distance bin (Figure 2.2a), by summing the size of an animal cluster rather than counting individual animals.

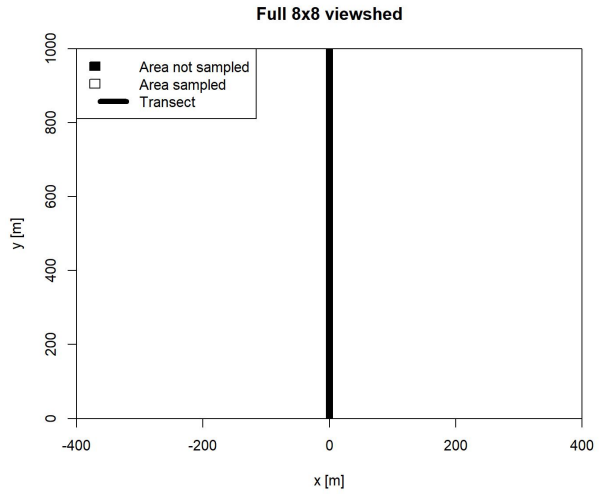
With the current implementation of traditional distance sampling it was not possible to use cluster sizes. Therefore, when an abundance estimate was done using cluster sizes, fitting the distance distribution of detected animals before it was viewshed compensated was used as baseline distance sampling (green distribution in Figure 2.2a). This alternative to traditional distance sampling is called baseline distance sampling. With a complete viewshed baseline distance sampling is very similar to traditional distance sampling. However this method fits the distribution of distance bins, whereas traditional distance sampling fits individual detections, resulting in a different abundance estimate.

2.5 Simulated input data

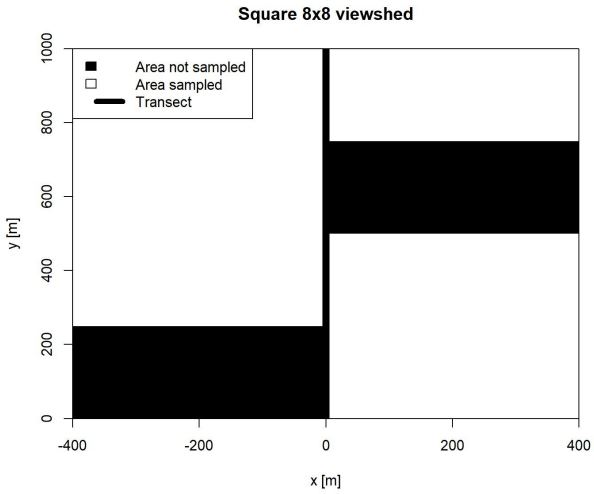
All methods discussed in this thesis required a raster with each cell being assigned a viewshed parameter, v. For Bayesian distance sampling the viewshed parameter indicates whether the area was sampled or not, with $v \in \{0, 1\}$. For other methods, the viewshed parameter describes the likelihood that an observer was able to see the animal from the transect, with $v \in [0, 1]$.

2.5.1 Simulated viewshed raster

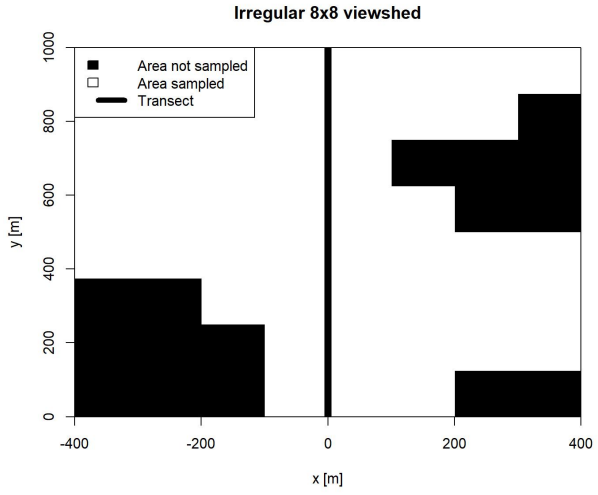
For Bayesian distance sampling, three viewshed rasters, with 8 by 8 cells, were used as defined in Figure 2.3. The half-width w is 400 metres and the transect length L is 1.000 metres. These are equivalent to the viewsheds used by (Delaney et al., 2025).



(a) Full viewshed



(b) Square viewshed

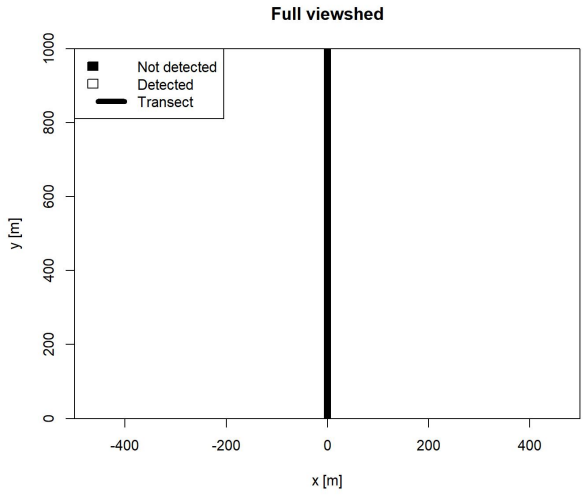


(c) Irregular viewshed

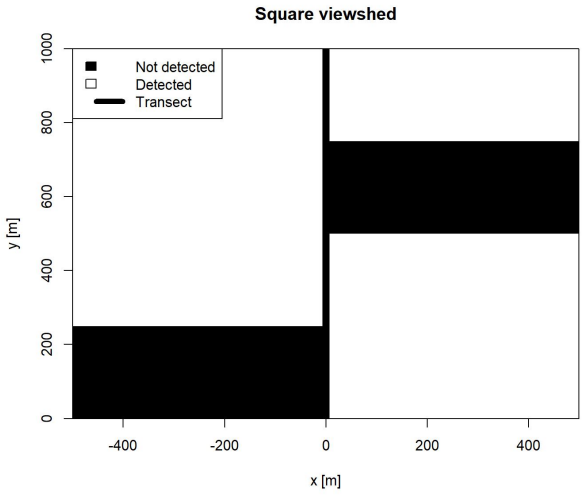
Figure 2.3: Manually created viewshed raster used for Bayesian distance sampling. Raster value indicates whether the area was sampled or not. 8x8 raster resolution. White, area sampled. Black, area not sampled. Transect at $x = 0$. $L = 1.000$, $w = 400$.

For Bayesian distance sampling the viewshed parameter is either 0 or 1, indicating whether the area was sampled or not. In the full viewshed (2.3a) all values are 1. This means that there is no influence of the viewshed on distance sampling. In the square viewshed (2.3b) two blocks of the viewshed are blocked off. For the methods discussed in this thesis, since animals are randomly distributed with respect to the observer (assumption 3), the sampled area functions as an area with a full viewshed but with both the area and population size reduced to 75%. The irregular viewshed (2.3c) is the most basic simulation of a real viewshed. It blocks various cells of the raster without a specific pattern.

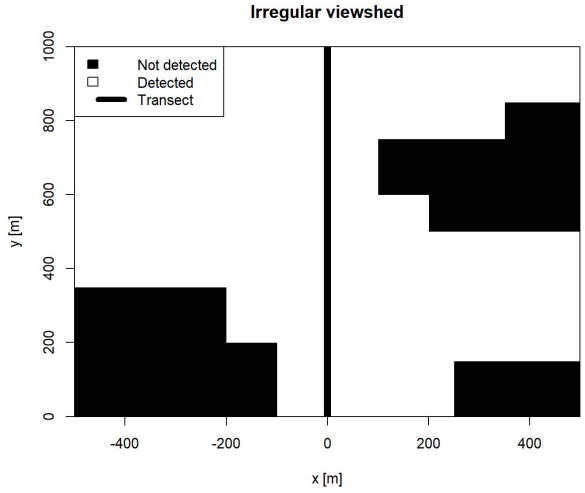
For my method, more viewsheds were designed. The resolution of the viewshed rasters was improved to 20 by 20 cells and values between 0 and 1 were allowed ($v \in [0, 1]$), rather than only values of 0 and 1 being allowed for Bayesian distance sampling ($v \in \{0, 1\}$). The half-width w is 500 metres and the transect length L is 1.000 metres, such that the area is 1 km^2. Viewsheds equivalent (or as close to equivalent) to the 8 by 8 viewsheds were developed. The full viewshed (Figure 2.4a) and square viewshed (Figure 2.4b) are equivalent. However the irregular viewshed (Figure 2.4c) could not be upsampled to the exact same shape. Therefore it has been made to have the same characteristics as the 8 by 8 irregular viewshed. Additionally a one sided viewshed (Figure 2.4d) was added, where the viewshed is only blocked on one side of the area. This viewshed was designed to expose biases in viewshed incorporated distance sampling when the viewshed is very different on both sides of the area. The complex viewshed (2.4e) is the closest to a simulation of a real viewshed. The raster values are now any value from 0, 0.1, 0.2, ..., 0.9, 1. The values have manually been applied by visual inspection. The logic is based on observers being located along the transect (in theory infinitely many). The value of a cell represents what percentage of observers can see the cell. A value of 1 means that the cell is visible for observers along the entire transect. A value of 0 indicates one of two options. It either represents an obstacle, or it cannot be seen by any observer along the transect due to obstacles. Any value between 0 and 1 indicates what percentage of observers can see the cell.



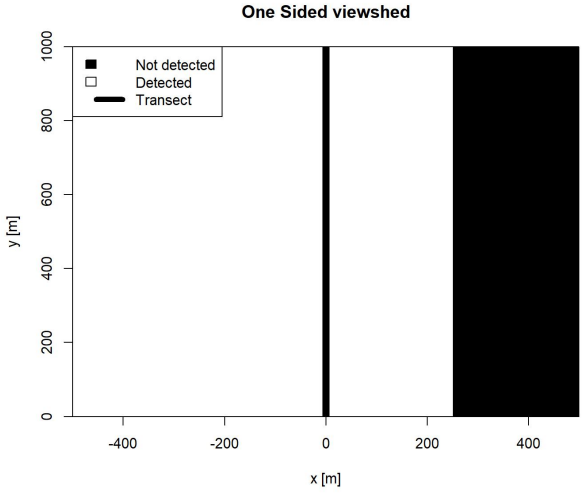
(a) Full viewshed



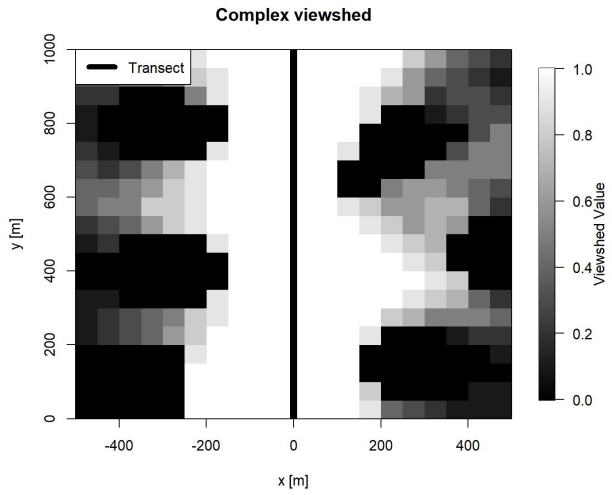
(b) Square viewshed



(c) Irregular viewshed



(d) One sided viewshed



(e) Complex viewshed

Figure 2.4: Manually created viewshed raster used for my method. Raster value indicates the probability of an animal being detected. 20x20 raster resolution. Transect at $x = 0$. $L = 1.000$, $w = 500$.

2.5.2 Simulated animal population (1)

The animal population is simulated according to the following steps (each step applies to any animal generated):

1. Generate location with Cartesian coordinates ($x \sim \text{Uniform}(0, L)$, $y \sim \text{Uniform}(-w, w)$).
2. Extract viewshed parameter by taking viewshed raster value at location.
3. Define distance to transect as the smallest Euclidean distance from the animal to the transect.
4. Generate animal cluster size, $s \sim \text{Uniform}(1, s_{max})$.
5. Determine if animal is detected based on distance to transect, by taking a binary sample, with probability $g(x)$ (equation 2.4).
6. Determine if the animal is detected based on the viewshed parameter, by taking a binary sample with the probability determined by the viewshed parameter.
7. Categorise as 'Detected' if both outputs of step 5 and 6 are true, otherwise categorise as 'Not detected'.

An example of a simulated animal population is shown in Figure 2.5.

The animals in the 'Detected' category are used as input for viewshed incorporated distance sampling (n). The true abundance is the number of animals in the total population (N). When cluster sizes are used, the true abundance increases by a factor of $\frac{1+s_{max}}{2}$.

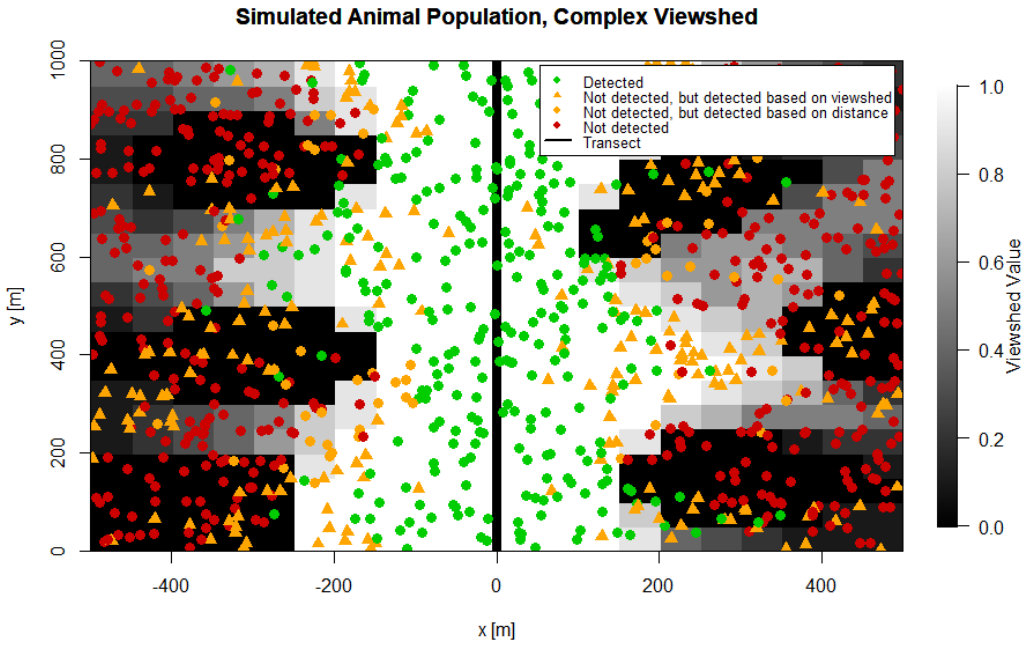


Figure 2.5: Example of simulated animal population, with complex viewshed underlayed. Green circle, detected animals. Red circle, undetected animals. Orange triangle, animals which would have been detected if detection only depends on viewshed parameter. Orange circle, animals which would have been detected if detection only depends on distance. Half-normal equation (equation 2.4) applied for distance detection. $L = 1.000$, $w = 500$.

2.6 Mixed input data

2.6.1 Study area

The study area is the southwestern part of the Kgalagadi Transfrontier Park (KTP, ∼38.000 km^2). The KTP originally consists of the Kalahari Gemsbok National Park (northwestern South Africa) and the Gemsbok National Park (southwestern Botswana) and straddles the southeastern border of Namibia (Moswete et al., 2012). The study area covers an area of ∼15.000 km^2 and is predominantly located within South Africa's boundaries, between 20°00' E and 21°3' E, and 24°33' S and 26°29' S, as shown in Figure 2.6.

The area is a semi-arid savanna, with two dry riverbeds, the Nossob and Auob rivers that meander through the area. Water has not flowed through these rivers since 1964 and 1974, respectively. However, there is still underground water flow which has been exploited by humans through the construction of boreholes for water extraction. These boreholes attract animal activity. Fences have been erected on the border between South Africa and Namibia due to farming activities on the Namibian side (Williamson and Williamson, 1984).

These activities, along with other anthropogenic activities, have impacted the migration of animals. Due to the fences being erected across migration routes, migration of wide ranging wildlife such as wildebeest and springbok has severely been restricted, while many species have disappeared entirely (Williamson and Williamson, 1984).

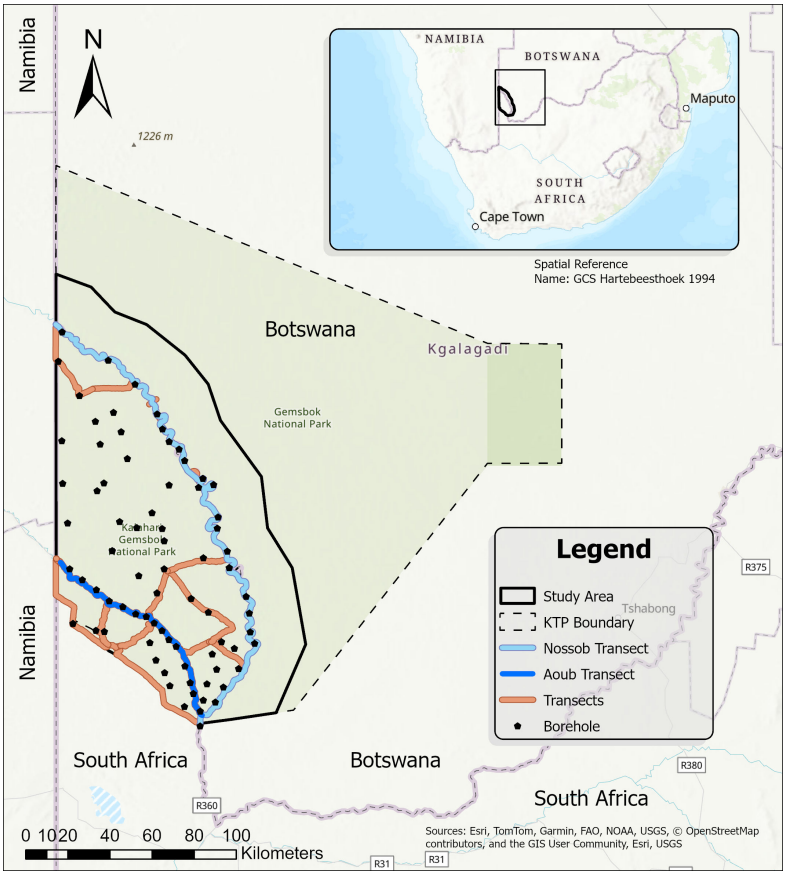


Figure 2.6: Study area in the Kgalagadi Transfrontier Park, South Africa. In orange, all transects driven. In light blue, the transect along the Aoub riverbed. In dark blue, the transect along the Nossob riverbed. Boreholes are indicated.

2.6.2 Real viewshed raster

To validate the simulated method a real case study of an applied viewshed from the KTP was used. Just like the simulated viewsheds these were defined as a raster. The Digital Elevation Model (DEM), used for this was provided by GeoSMART. The DEM had a resolution of 5 by 5 metres, and used the Hartebeesthoek94 Geographic Coordinate System (EPSG, 2026). First, the transects along which to calculate the viewshed had to be defined. These transects are polylines of the routes which were driven. The polyline transect was based on GPS tracks which were recorded every 10 seconds during driving. If these were not available, a combination of observer locations, a polyline of transects driven for earlier research and roads displayed on the background map, were used to manually draw the transect polyline.

Along this transect polyline, points were placed at regular intervals. These points were used as observer locations for the viewshed calculation. The viewshed was calculated using the Geodesic Viewshed tool in ArcGIS (ArcGIS, 2026). The viewshed was calculated by counting how many of the observer locations have line of sight to a raster cell based on the DEM. Before the viewshed calculation, a vertical offset of 1 metre was applied to represent the observer's eye height above the ground. The maximum viewing distance from the transect to calculate the viewshed was set at 350 metres. 350 metres was chosen as during game counts in the KTP it was found that for larger distances, measurements were inaccurate. For a straight transect there is a maximum viewshed value for a raster cell, which is defined by the observer location interval and the maximum viewing distance. If the transect is not straight it is possible to exceed this maximum viewshed value. The total output of the viewshed calculation is the count of how many observer locations were able to see the cell. For applications into viewshed incorporated methods it is divided by the straight transect maximum viewshed value of 21.

To validate the method a subsample of the study area was used. This subsample was chosen to be the Nossob riverbed transect, running along the entire length of the Nossob riverbed (Figure 2.6). This was done for multiple reasons. (1) Animals are more abundant in the riverbeds than the dune areas, so more animals were counted in the riverbed. This allows for more accurate abundance estimation. (2) Some parts of the Nossob riverbed have been sampled multiple times, which allows for more in-depth analysis, where either an abundance estimate can be made with more data or multiple abundance estimates can be compared. (3) The computation time for the viewshed was manageable (4 hours). From here on, this viewshed will be called the Nossob viewshed.

2.6.3 Simulated animal population (2)

For a real viewshed, the animal population is simulated according to the following steps (each step applies to any animal generated):

1. Randomly create location with x-coordinate and y-coordinate within the extent of the viewshed raster. In this step the viewshed parameter is assigned to each animal.

2. Normalise the viewshed parameter according to the maximum viewshed parameter along a straight line.

3. Define distance to transect as the smallest Euclidean distance from the animal to the transect.

4. Generate animal cluster size, $s \sim \mathrm{Uniform}(1, s_{max})$.

5. Determine if animal is detected based on distance to transect, by taking a binary sample, with probability $g(x)$ (equation 2.4).

6. Determine if the animal is detected based on the viewshed parameter, by taking a binary sample with the probability determined by the viewshed parameter.

7. Categorise as 'Detected' if both outputs of step 5 and 6 are true, otherwise categorise as 'Not detected'.

This method results in an equivalent simulated animal population to the animal population simulation method outlined in section 2.5.2, with the difference being the extent of possible animal location and how the viewshed parameter is assigned.

2.7 Real input data

2.7.1 Real animal population

To evaluate the performance of viewshed incorporated distance sampling, game counts (systematic method to determine wildlife abundance in a predefined area (Pacey, 2026)) obtained in the study area were used. The recorded attributes which are important for evaluating the performance of viewshed incorporated distance sampling are displayed in Table 2.1. More detailed information on the game counts can be found in Appendix A.2.

Table 2.1: Explanation of the attributes in the validation dataset.

Attribute	**Explanation**
Distance	Distance to centre of cluster
Bearing	Angle relative to north
Latitude	Latitude of observer
Longitude	Longitude of observer
Animal count adult male	Number of adult male animals in the cluster
Animal count adult female	Number of adult female animals in the cluster
Animal count juvenile	Number of juvenile animals in the cluster
Animal count calf	Number of calf animals in the cluster
Animal count unknown	Number of unknown animals in the cluster

For the real animal population, the game counts validation data was filtered to only contain the game counts performed along the Nossob riverbed. The animal counts of different categories were summed to a total animal count. Next the location of animals was calculated. The x- and y-location of the observer, the distance to the centre of the animal cluster and the bearing to the centre of the animal cluster were used to determine the x- and y-location of animals. The resulting animal population is equivalent to the animal population categorised as 'Detected' in section 2.5.2 and section 2.6.3. An example of the resulting animal locations is shown in Figure 2.7.

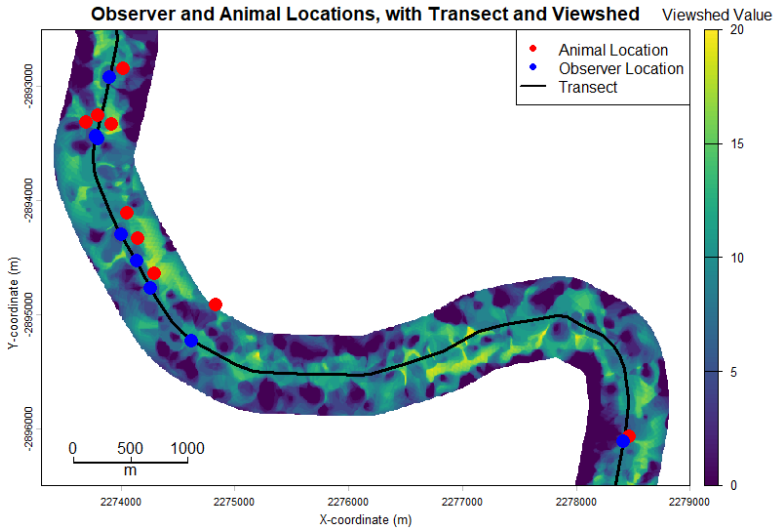


Figure 2.7: Example of processed game counts with the Nossob viewshed. In red, animal locations. In blue, observer locations. In black, the transect. Some observer location overlap, thus looking like one observer location. Maximum viewshed value along a straight transect is 21, with the maximum viewshed value in this example being 20.

2.8 Performance evaluation

To evaluate the performance of methods using simulated input data and mixed input data, a standardised method was used.

To compare results with different animal abundances, the abundance estimate bias was calculated by normalising the abundance estimate to the true abundance:

$$\hat{N}_{bias} = \frac{\hat{N} - N}{N} \tag{2.13}$$

An abundance estimate bias of 0 means that the abundance estimate is equal to the true abundance. The abundance estimate bias can also be expressed as a percentage.

Each distance sampling method produced an abundance estimate bias. Multiple simulations of the method were run, resulting in a distribution of abundance estimate biases. The mean, median and standard deviation of the abundance estimate bias distribution were calculated. Unless otherwise indicated, any reference to mean, median or standard deviation concerns the mean, median or standard deviation of the abundance estimate bias distribution. Additionally, the mean value of the last 10% of the abundance estimate bias distribution was calculated.

To evaluate convergence of the abundance estimate bias distribution across multiple simulations, the cumulative mean of the abundance estimate bias was calculated:

$$\bar{\hat{N}}_{bias} = \frac{1}{m} \sum_{i=1}^{m} \hat{N}_{bias} \tag{2.14}$$

Convergence is evaluated based on visual inspection of the progression of the cumulative mean of the abundance estimate bias and the similarity between the mean of the last 10% and the overall mean. If visually, the progression of the cumulative mean of the abundance estimate bias has stabilised, and the mean of the last 10% of the abundance estimate bias distribution is similar to the mean of the

full abundance estimates bias distribution, convergence is achieved.

Performance is based on the mean, median and standard deviation of the abundance estimate bias distribution, and visual inspection of the combination of a violin plot overlaid with a boxplot. High performance is achieved with a mean and median close to 0 and a small standard deviation, and the violin plot showing a normal distribution. All performance evaluation concerning mean, median, and standard deviation, is relative to the mean, median and standard deviation of other abundance estimate bias distribution.

3 Results

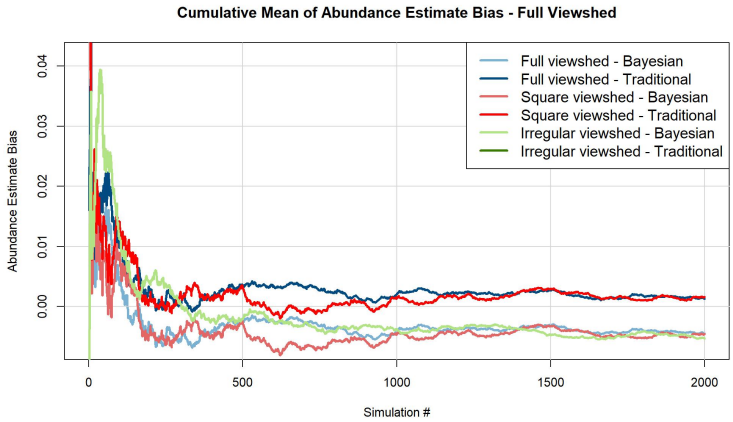
3.1 Bayesian distance sampling

For Bayesian distance sampling the true abundance N is not equal to the abundance used for the abundance estimate bias $\hat{N}_{bias}$ calculation (equation 2.13) for incomplete viewsheds. For the square viewshed and irregular viewshed, the sampled area is 75% of the true area. Therefore, the abundance used to calculate the abundance estimate bias is 75% of the true abundance. This is consistent with the implementation in Delaney et al. (2025).

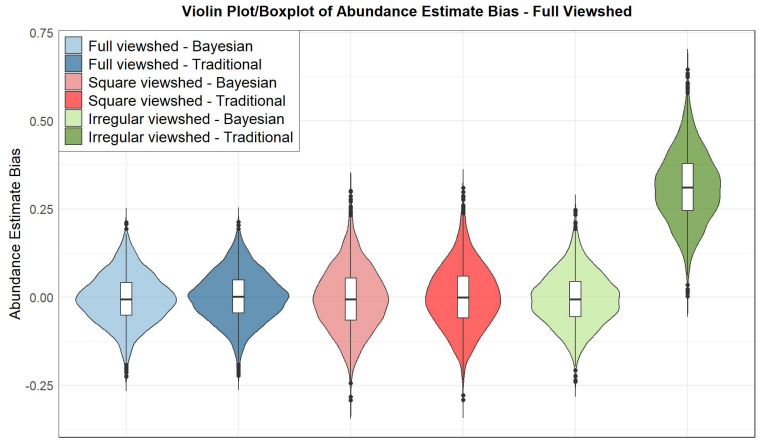
Bayesian distance sampling is compared to traditional distance sampling, for the full viewshed, square viewshed and irregular viewshed, in Figure 3.1 and Table 3.1.

Table 3.1: Bayesian distance sampling various viewshed comparison. Methods applied: Bayesian distance sampling and traditional distance sampling. No cluster sizes used. Number of simulations = 2.000, population size = 500. Input $\sigma = 150$.

Method	mean	median	sd	mean of last 10%
Full viewshed, Bayesian	-0.00448	-0.00590	0.06968	-0.00420
Full viewshed, traditional	0.00128	0.00136	0.06920	0.00160
Square viewshed, Bayesian	-0.00453	-0.00650	0.08879	-0.00475
Square viewshed, traditional	0.00163	-0.00088	0.08819	0.00142
Irregular viewshed, Bayesian	-0.00532	-0.00661	0.07359	-0.00473
Irregular viewshed, traditional	0.31256	0.31072	0.09949	0.31337



(a) Cumulative mean of abundance estimate bias. Traditional distance sampling outside plot extent.



(b) Violin plot and boxplot of abundance estimate bias.

Figure 3.1: Bayesian distance sampling various viewshed comparison. Methods applied: Bayesian distance sampling and traditional distance sampling. No cluster sizes used. Number of simulations = 2.000, population size = 500. Input $\sigma = 150$.

For all viewsheds and both distance sampling methods in Figure 3.1 and Table 3.1, the cumulative mean of the abundance estimate bias has converged. The performance of Bayesian distance sampling is similar for all three viewsheds. For the full viewshed and the square viewshed, the performance of Bayesian distance sampling is worse than the performance of traditional distance sampling. For the irregular viewshed, Bayesian distance sampling outperforms traditional distance sampling.

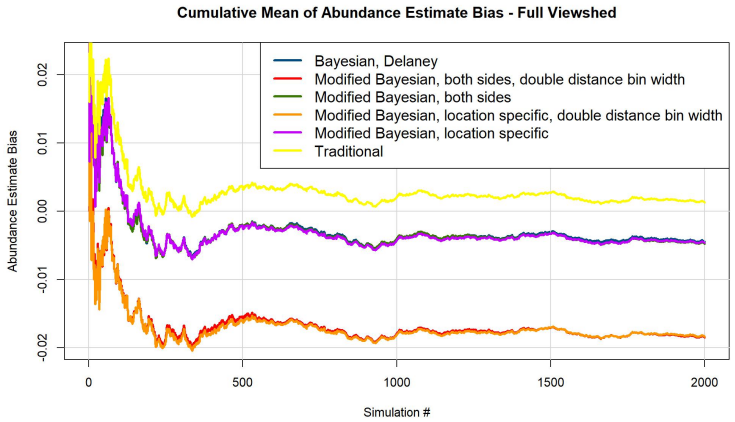
The performance of Bayesian distance sampling on the irregular viewshed is similar to the performance of Bayesian distance sampling on the full viewshed and square viewshed, indicating that the method performs similarly irrespective of the viewshed used.

3.2 Modified Bayesian distance sampling

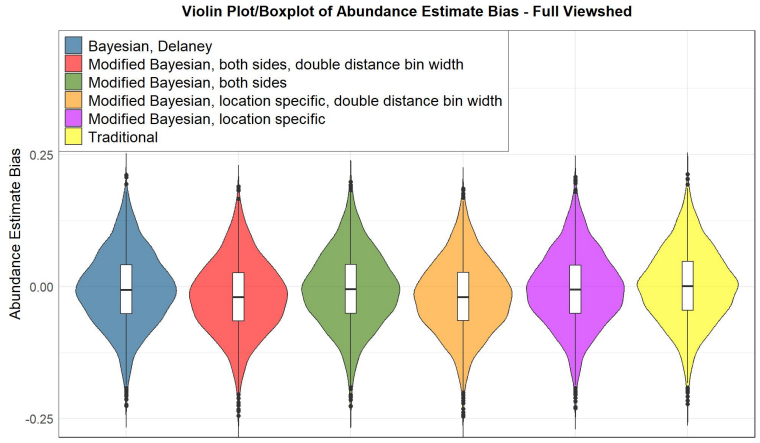
Modified Bayesian distance sampling is compared to Bayesian distance sampling and traditional distance sampling, for the full viewshed, in Figure 3.2 and Table 3.2.

Table 3.2: Various Bayesian distance sampling methods comparison. Methods applied: Bayesian distance sampling by Delaney et al. (2025), modified Bayesian distance sampling and traditional distance sampling. Full viewshed. No cluster sizes used. Number of simulations = 2.000, population size = 500. Input $\sigma = 150$.

Method	mean	median	sd	mean of last 10%
Bayesian, Delaney	-0.00448	-0.00590	0.06968	-0.00420
Modified Bayesian, both sides, double distance bin width	-0.01849	-0.01966	0.06865	-0.01820
Modified Bayesian, both sides,	-0.00478	-0.00475	0.06971	-0.00449
Modified Bayesian, location specific, double distance bin width	-0.01839	-0.01942	0.06864	-0.01812
Modified Bayesian, location specific,	-0.00466	-0.00542	0.06970	-0.00441
Traditional	0.00128	0.00136	0.06920	0.00160



(a) Cumulative mean of abundance estimate bias.



(b) Violin plot and boxplot of abundance estimate bias.

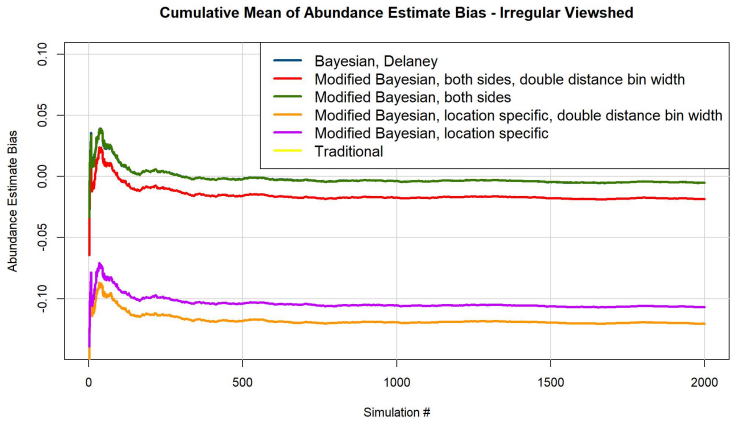
Figure 3.2: Various Bayesian distance sampling methods comparison. Methods applied: Bayesian distance sampling by Delaney et al. (2025), modified Bayesian distance sampling and traditional distance sampling. Full viewshed. No cluster sizes used. Number of simulations = 2.000, population size = 500. Input $\sigma = 150$.

For all methods in Figure 3.2 and Table 3.2, the cumulative mean of the abundance estimate bias has converged. The performance for modified Bayesian distance sampling with the same distance bin width used by Delaney et al. (2025), is similar to the performance of Bayesian distance sampling by Delaney et al. (2025). However, the performance of modified Bayesian distance sampling with double distance bin width is worse than the performance of Bayesian distance sampling by Delaney et al. (2025). This indicates that the distance bin width is crucial to the performance of Bayesian distance sampling.

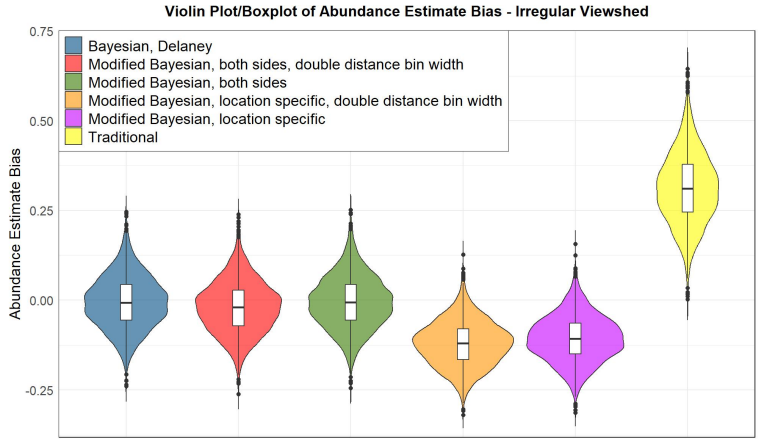
The modified versions of Bayesian distance sampling are compared to the implementation of Bayesian distance sampling of Delaney et al. (2025) and to traditional distance sampling, for the irregular viewshed, in Figure 3.3 and Table 3.3.

Table 3.3: Various Bayesian distance sampling methods comparison. Methods applied: Bayesian distance sampling by Delaney et al. (2025), modified Bayesian distance sampling and traditional distance sampling. Irregular viewshed. No cluster sizes used. Number of simulations = 2.000, population size = 500. Input $\sigma = 150$.

Method	mean	median	sd	mean of last 10%
Bayesian, Delaney	-0.00532	-0.00661	0.07359	-0.00473
Modified Bayesian, both sides, double distance bin width	-0.01869	-0.01917	0.07294	-0.01803
Modified Bayesian, both sides,	-0.00536	-0.00572	0.07354	-0.00475
Modified Bayesian, location specific, double distance bin width	-0.12054	-0.12002	0.06304	-0.11995
Modified Bayesian, location specific,	-0.10694	-0.10724	0.06407	-0.10645
Traditional	0.31256	0.31072	0.09949	0.31337



(a) Cumulative mean of abundance estimate bias. Traditional distance sampling out of plot extent.



(b) Violin plot and boxplot of abundance estimate bias.

Figure 3.3: Various Bayesian distance sampling methods comparison. Methods applied: Bayesian distance sampling by Delaney et al. (2025), modified Bayesian distance sampling and traditional distance sampling. Irregular viewshed. No cluster sizes used. Number of simulations = 2.000, population size = 500. Input $\sigma = 150$.

For all methods in Figure 3.3 and Table 3.3, the cumulative mean of the abundance estimate bias has converged. The performance of both Bayesian distance sampling and modified Bayesian distance sampling is better than the performance of traditional distance sampling.

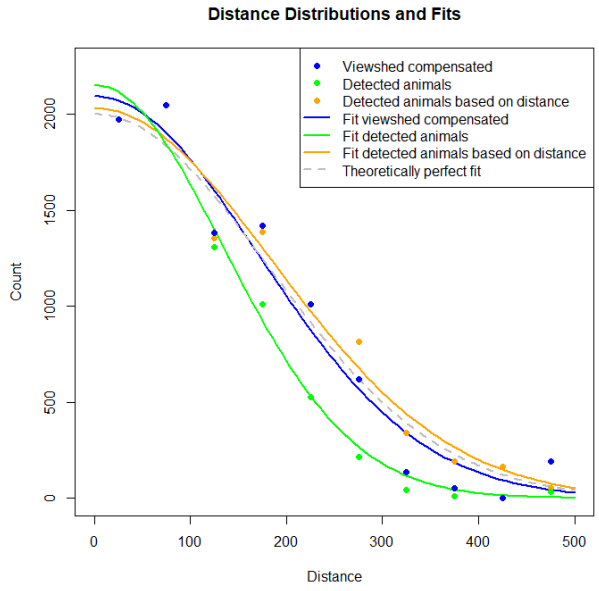
For modified Bayesian distance sampling using distance bins on both sides of the transect, performance is similar to Bayesian distance sampling by Delaney et al. (2025) when using the same distance bin width. With performance being worse for modified Bayesian distance sampling using distance bins on both sides of the transect when double distance bin width is used. This is consistent with the results for the full viewshed.

However, the results are not consistent with the results for the full viewshed for location specific modified Bayesian distance sampling. With both distance bin widths, performance is worse than the performance for Bayesian distance sampling by Delaney et al. (2025). This indicates that modified Bayesian distance sampling at best has similar performance to Bayesian distance sampling by Delaney et al. (2025), but in many cases has worse performance.

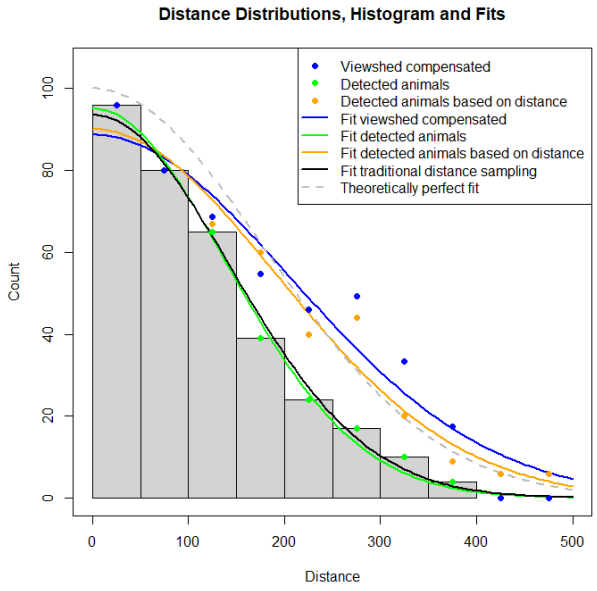
3.3 Viewshed compensated distance sampling

3.3.1 Simulated input data

In Figure 3.4 an example is shown of viewshed compensated distance sampling applied to the complex viewshed. Both with and without cluster sizes used.



(a) Cluster sizes used. For viewshed compensated distance distribution (blue): $\sigma = 170.52$, $\tau = 2093.59$. For detected animals distance distribution (green): $\sigma = 134.66$, $\tau = 2151.52$. For detected animals based on distance distribution (orange), $\sigma = 185.73$, $\tau = 2030.15$. Theoretical population size = 20.000.



(b) No cluster sizes used. Histogram of distances of detected animals. For viewshed compensated distance distribution (blue): $\sigma = 206.27$, $\tau = 88.72$. For detected animals distance distribution (green): $\sigma = 138.74$, $\tau = 95.19$. For detected animals based on distance distribution (orange), $\sigma = 191.17$, $\tau = 90.11$. In black, traditional distance sampling fit, $\sigma = 142.80$, $\tau = 93.63$. Population size = 1.000.

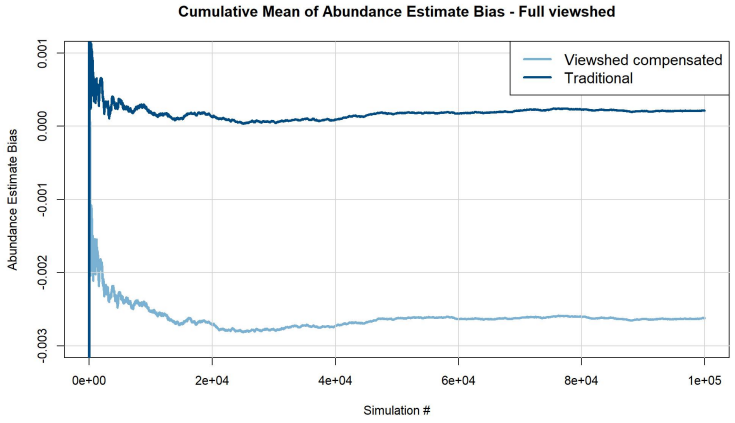
Figure 3.4: Overview of results of viewshed compensated distance sampling. Various distance distributions and their fits. Comparison between using cluster sizes and not using cluster sizes. In blue, the viewshed compensated distance distribution and fit. In green, the distribution of detected animals and fit. In orange, the distribution of animals which were not detected, but would have been detected if detection was only based on distance and the fit. In grey (dashed), the theoretically perfect, $\sigma = 180$, $\tau = 100$. Incomplete viewshed. Data points at low distance overlap.

In Figure 3.4 the improvement from fitting the uncompensated distance distribution to fitting the distance distribution compensated with the viewshed parameter distribution, both with cluster sizes used and without cluster sizes used, can be seen. As the viewshed and animal population were simulated, the theoretically perfect fit (dashed grey) and the fit of animals which would have been detected if detection was only based on distance are available (orange).

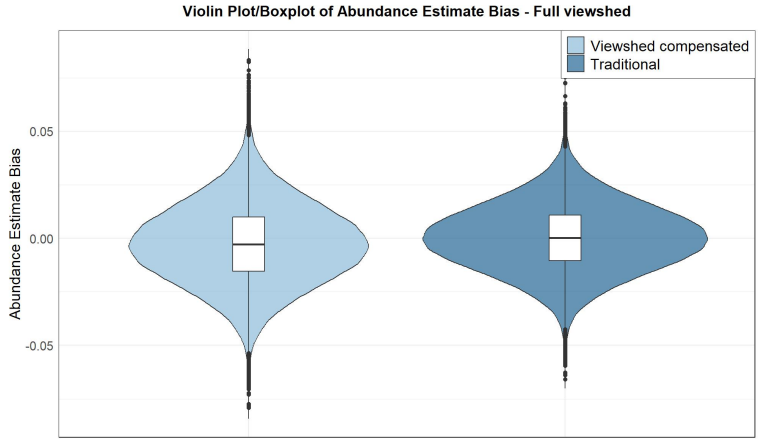
Viewshed compensated distance sampling is compared to traditional distance sampling, for the full viewshed, in Figure 3.5 and Table 3.4.

Table 3.4: Comparison viewshed compensated distance sampling and traditional distance sampling. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. Full viewshed. No cluster sizes used. Number of simulations = 100.000, population size = 10.000. Input $\sigma = 180$.

Method	mean	median	sd	mean of last 10%
Viewshed compensated	-0.00262	-0.00284	0.018892	-0.00263
Traditional	0.00021	0.00013	0.01582	0.00021



(a) Cumulative mean of abundance estimate bias.



(b) Violin plot and boxplot of abundance estimate bias.

Figure 3.5: Comparison viewshed compensated distance sampling and traditional distance sampling. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. Full viewshed. No cluster sizes used. Number of simulations = 100.000, population size = 10.000. Input $\sigma = 180$.

For both methods in Figure 3.5 and Table 3.4, the cumulative mean of the abundance estimate bias has converged.

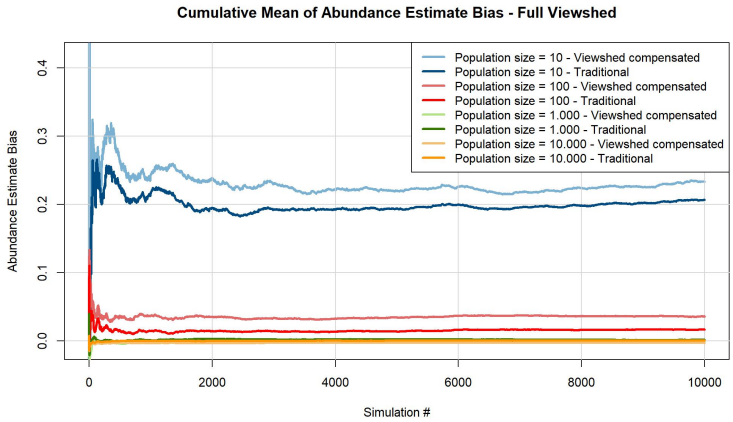
The performance of traditional distance sampling is better than the performance of viewshed compensated distance sampling. However, for both methods the mean and median are close to 0,

indicating good performance.

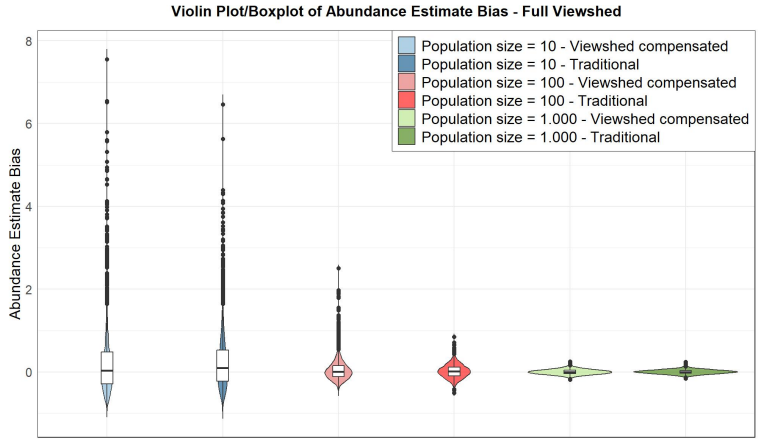
Viewshed compensated distance sampling is compared to traditional distance sampling, for the full viewshed, with different population sizes, in Figure 3.6 and Table 3.5.

Table 3.5: Comparison different population sizes. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. Full viewshed. No cluster sizes used. Number of simulations = 10.000. Input $\sigma = 180$.

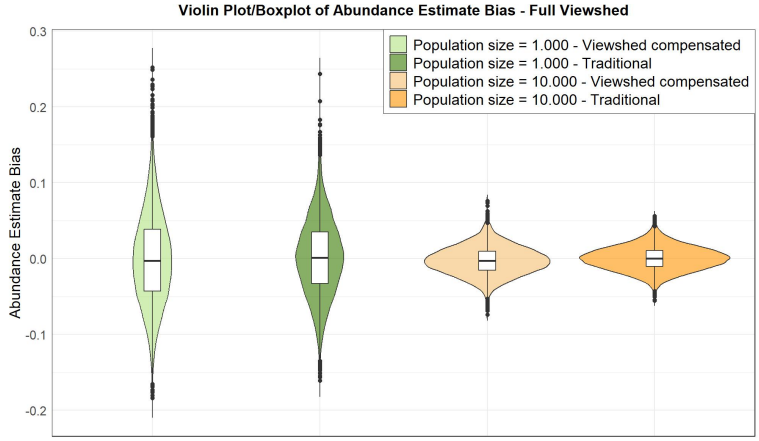
Method	mean	median	sd	mean of last 10%
Population size = 10, viewshed compensated	0.23311	0.03151	0.82647	0.23012
Population size = 10, traditional	0.20649	0.09909	0.60820	0.20440
Population size = 100, traditional	0.01651	0.01130	0.15935	0.01633
Population size = 100, viewshed compensated	0.03562	0.00630	0.22378	0.03574
Population size = 1.000, traditional	0.00124	0.00103	0.05015	0.00115
Population size = 1.000, viewshed compensated	-0.00065	-0.00276	0.06028	-0.00085
Population size = 10.000, traditional	0.00013	0.00019	0.01566	0.00012
Population size = 10.000, viewshed compensated	-0.00284	-0.00302	0.01862	-0.00284



(a) Cumulative mean of abundance estimate bias.



(b) Violin plot and boxplot of abundance estimate bias. Population sizes 10, 100 and 1.000.



(c) Violin plot and boxplot of abundance estimate bias. Population sizes 1.000 and 10.000.

Figure 3.6: Comparison different population sizes. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. No cluster sizes used. Full viewshed. Number of simulations = 10.000. Input $\sigma = 180$.

For all methods in Figure 3.6 and Table 3.5, the cumulative mean of the abundance estimate bias has converged.

A larger population size leads to faster convergence and better performance. The median is affected less by a larger population size as the mean. From this it could be concluded that outliers

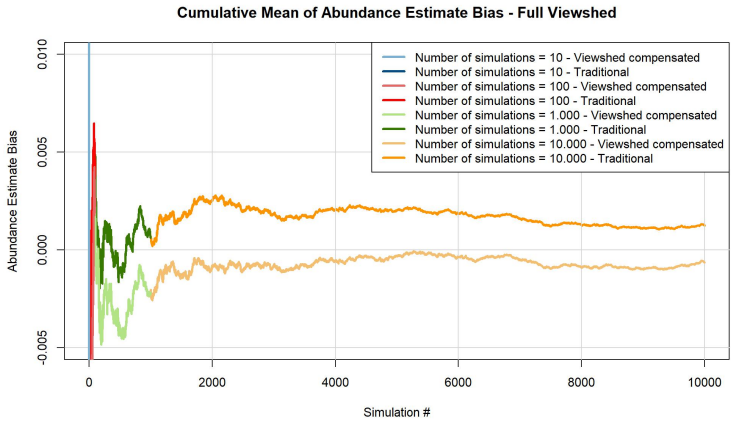
tend to have a large positive abundance estimate bias, leading to a higher positive mean while the median is still close to 0.

For traditional distance sampling, every step of a larger population size leads to better performance. For viewshed compensated distance sampling the same trend is observed. However, going from population size 1.000 to 10.000, the mean and median go further away from 0, while the standard deviation still decreases. This indicates an inherent bias in the viewshed compensated distance sampling method. Whereas for increasing population sizes the mean and median for traditional distance sampling converge to 0, the mean and median for viewshed compensated distance sampling converge to a non-zero value.

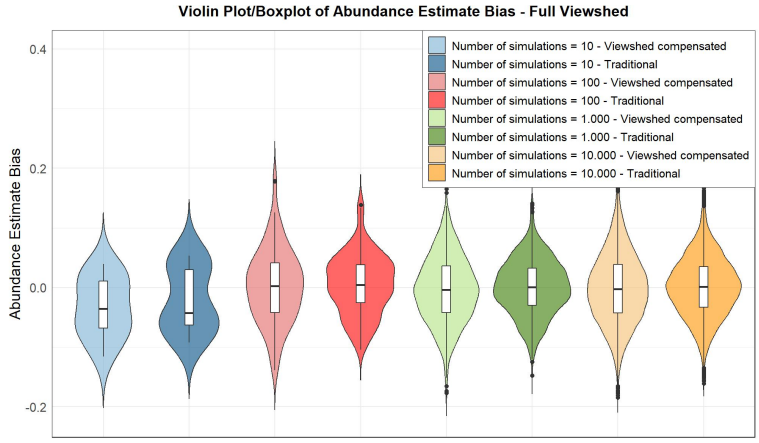
Viewshed compensated distance sampling is compared to traditional distance sampling, for the full viewshed, with a different number of simulations, in Figure 3.7 and Table 3.6.

Table 3.6: Comparison of different number of simulations. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. Full viewshed. No cluster sizes used. Population size = 1.000. Input $\sigma = 180$.

Method	mean	median	sd	mean of last 10%
Number of simulations = 10, viewshed compensated	-0.03294	-0.03546	0.05091	-0.03294
Number of simulations = 10, traditional	-0.02117	-0.04248	0.05544	-0.02117
Number of simulations = 100, viewshed compensated	0.00265	0.00232	0.06260	0.00242
Number of simulations = 100, traditional	0.00462	0.00464	0.05088	0.00419
Number of simulations = 1.000, viewshed compensated	-0.00229	-0.00421	0.05781	-0.00208
Number of simulations = 1.000, traditional	0.00062	0.00085	0.04699	0.00090
Number of simulations = 10.000, viewshed compensated	-0.00065	-0.00276	0.06028	-0.00085
Number of simulations = 10.000, traditional	0.00124	0.00103	0.05015	0.00115



(a) Cumulative mean of abundance estimate bias.



(b) Violin plot and boxplot of abundance estimate bias.

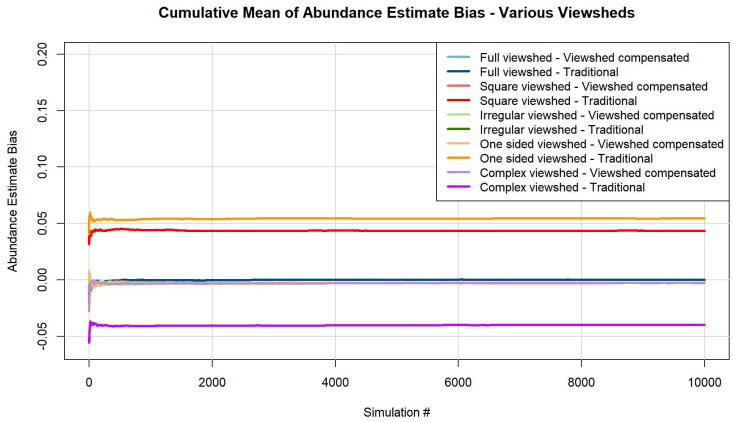
Figure 3.7: Comparison of different number of simulations. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. No cluster sizes used. Full viewshed. Population size = 1.000. Input $\sigma = 180$.

In Figure 3.7 and Table 3.6 it can be seen that for both viewshed compensated distance sampling and traditional distance sampling, the performance increases with more simulations. The cumulative mean of the abundance estimate bias has converged around 3500 simulations. Therefore, the performance for 1.000 simulations being better than the performance with 10.000 simulations, is a result of the randomisation in the simulations.

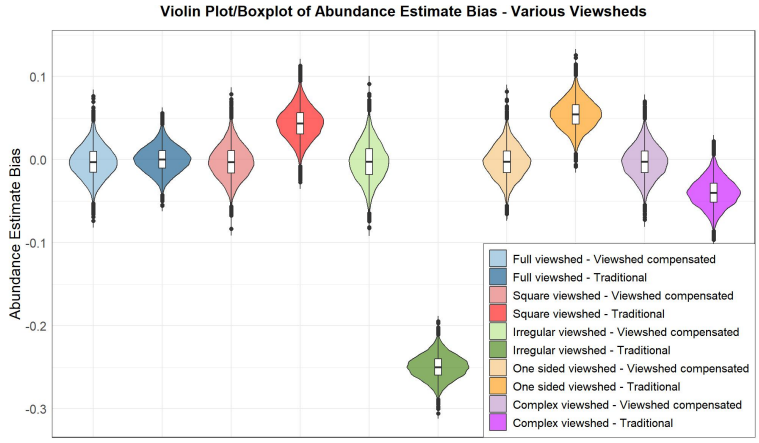
Viewshed compensated distance sampling is compared to traditional distance sampling, for the full viewshed, the square viewshed, the irregular viewshed, the one sided viewshed and the complex viewshed, in Figure 3.8 and Table 3.7.

Table 3.7: Various viewsheds comparison. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. No cluster sizes used. Number of simulations = 10.000, population size = 10.000. Input $\sigma = 180$.

Method	mean	median	sd	mean of last 10%
Full viewshed, viewshed compensated	-0.00284	-0.00302	0.01862	-0.00284
Full viewshed, traditional	0.00013	0.00019	0.01566	0.00012
Square viewshed, viewshed compensated	-0.00277	-0.00312	0.01967	-0.00281
Square viewshed, traditional	0.04349	0.04366	0.01910	0.04347
Irregular viewshed, viewshed compensated	-0.00270	-0.00283	0.02248	-0.00273
Irregular viewshed, traditional	-0.24993	-0.24989	0.01434	-0.24997
One sided viewshed, viewshed compensated	-0.00278	-0.00288	0.01910	-0.00279
One sided viewshed, traditional	0.05430	0.05432	0.01731	0.05429
Complex viewshed, viewshed compensated	-0.00263	-0.00263	0.01932	-0.00268
Complex viewshed, traditional	-0.04004	-0.04008	0.01711	-0.04008



(a) Cumulative mean of abundance estimate bias. Irregular viewshed - Viewshed compensated distance sampling outside plot extent.



(b) Violin plot and boxplot of abundance estimate bias.

Figure 3.8: Various viewsheds comparison. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. No cluster sizes used. Number of simulations = 10.000, population size = 10.000. Input $\sigma = 180$.

For all methods in Figure 3.8 and Table 3.7, the cumulative mean of the abundance estimate bias has converged.

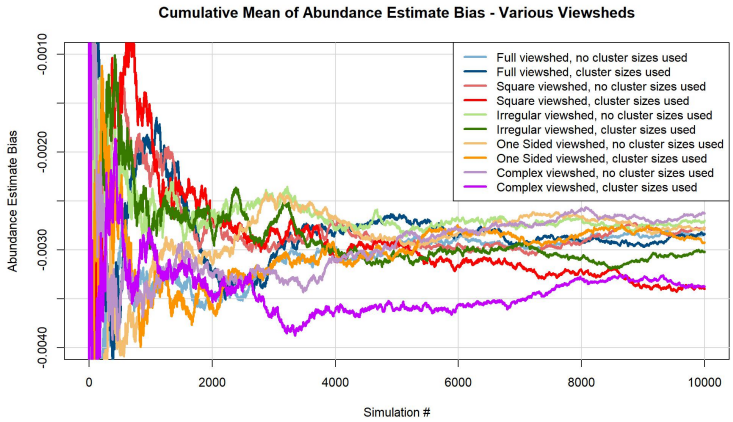
For viewshed compensated distance sampling across all viewsheds the performance is very similar, while for traditional distance sampling, incomplete viewsheds cause a large drop in performance compared to the full viewshed. This indicates that the viewshed does not impact performance for viewshed compensated distance sampling, whereas the viewshed does impact performance for traditional distance sampling.

Given the means and standard deviations for the complex viewshed in Table 3.7 and the population size of 10.000, the abundance estimate with 95% confidence interval for viewshed compensated distance sampling is $\hat{N} = 9974 \pm 741$ and the abundance estimate with 95% confidence interval for traditional distance sampling is $\hat{N} = 9600 \pm 631$.

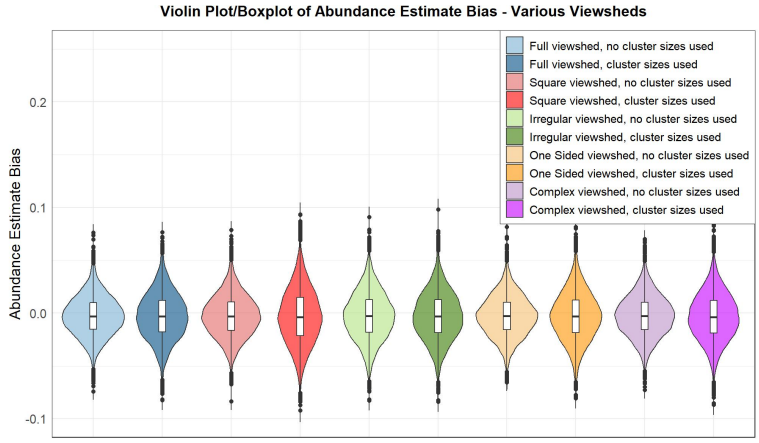
Viewshed compensated distance sampling with cluster sizes is compared to viewshed compensated distance sampling without cluster sizes, for the full viewshed, the square viewshed, the irregular viewshed, the one sided viewshed and the complex viewshed, in Figure 3.9 and Table 3.8.

Table 3.8: Comparison viewshed compensated distance sampling with and without animal cluster sizes used with various viewsheds. Method applied: Viewshed compensated distance sampling. Number of simulations = 10.000, population size = 10.000. Input $\sigma = 180$.

Method	mean	median	sd	mean of last 10%
Full viewshed, no cluster sizes used	-0.00284	-0.00302	0.01862	-0.00284
Full viewshed, cluster sizes used	-0.00284	-0.00318	0.02208	-0.00288
Square viewshed, no cluster sizes used	-0.00277	-0.00312	0.01967	-0.00281
Square viewshed, cluster sizes used	-0.00340	-0.00376	0.02616	-0.00339
Irregular viewshed, no cluster sizes used	-0.00270	-0.00283	0.02248	-0.00273
Irregular viewshed, cluster sizes used	-0.00302	-0.00329	0.02267	-0.00305
One sided viewshed, no cluster sizes used	-0.00278	-0.00288	0.01910	-0.00279
One sided viewshed, cluster sizes used	-0.00293	-0.00307	0.02291	-0.00285
Complex viewshed, no cluster sizes used	-0.00263	-0.00263	0.01932	-0.00268
Complex viewshed, cluster sizes used	-0.00338	-0.00371	0.02245	-0.00334



(a) Cumulative mean of abundance estimate bias.



(b) Violin plot and boxplot of abundance estimate bias.

Figure 3.9: Comparison viewshed compensated distance sampling with and without animal cluster sizes used with various viewsheds. Method applied: Viewshed compensated distance sampling. Number of simulations = 10.000, population size = 10.000. Input $\sigma = 180$.

For all methods in Figure 3.9 and Table 3.8, the cumulative mean of the abundance estimate bias has converged.

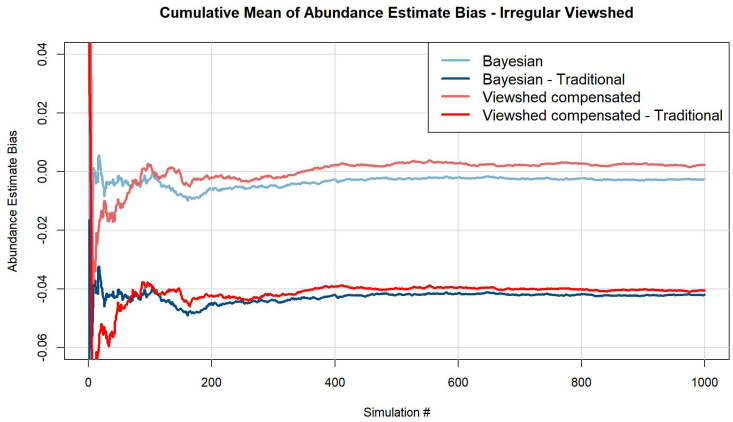
The performance with cluster sizes used is worse than the performance without cluster sizes used across all viewsheds. The mean and median being further away from 0 indicates that the inherent bias of viewshed compensated distance sampling increases when cluster sizes are applied. The larger standard deviation is explained by the extra randomness that is introduced by enabling cluster sizes.

Viewshed compensated distance sampling without cluster sizes used is compared to Bayesian distance sampling and traditional distance sampling with the irregular viewshed, in Figure 3.10 and Table 3.9.

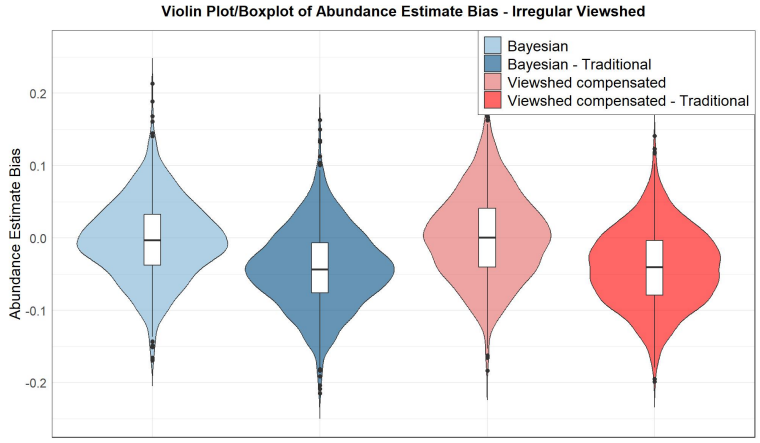
There are 2 differences from the implementation of Bayesian distance sampling used for the results in section 3.1. (1) The irregular viewshed with 20x20 cells was used (Figure 2.4c) instead of the irregular viewshed with 8x8 cells (Figure 2.3c). (2) The abundance estimate for Bayesian distance sampling was extrapolated to the full area.

Table 3.9: Comparison viewshed compensated distance sampling to Bayesian distance sampling and traditional distance sampling. Methods applied: Viewshed compensated distance sampling, Bayesian distance sampling and traditional distance sampling. Irregular viewshed 20x20 cells. No cluster sizes used. Number of simulations = 10.000, population size = 10.000. Input $\sigma = 180$.

Method	mean	median	sd	mean of last 10%
Bayesian	-0.00258	-0.00307	0.05435	-0.00274
Traditional, ran with Bayesian	-0.04193	-0.04346	0.05475	-0.04210
Viewshed compensated	0.00222	0.00030	0.06026	0.00218
Traditional, ran with viewshed compensated	-0.04052	-0.04019	0.05287	-0.04053



(a) Cumulative mean of abundance estimate bias. Traditional distance sampling outside plot extent.



(b) Violin plot and boxplot of abundance estimate bias.

Figure 3.10: Comparison viewshed compensated distance sampling to Bayesian distance sampling and traditional distance sampling. Methods applied: Viewshed compensated distance sampling, Bayesian distance sampling and traditional distance sampling. Irregular viewshed 20x20 cells. No cluster sizes used. Number of simulations = 10.000, population size = 10.000. Input $\sigma = 180$.

For all methods in Figure 3.10 and Table 3.9, the cumulative mean of the abundance estimate bias has converged.

Viewshed compensated distance sampling and Bayesian distance sampling perform similarly. Traditional distance sampling run with the different methods does not perform equally for both methods, because the input data differed between methods.

Given the means and standard deviations in Table 3.9 and the population size of 10.000, the abundance estimate with 95% confidence interval for viewshed compensated distance sampling is $\hat{N} = 10022 \pm 2321$ and the abundance estimate with 95% confidence interval for Bayesian distance sampling is $\hat{N} = 9974 \pm 2084$.

3.3.2 Mixed input data

In Figure 3.11 a part of the Nossob viewshed is shown. Circular artefacts are seen in some parts of the area.

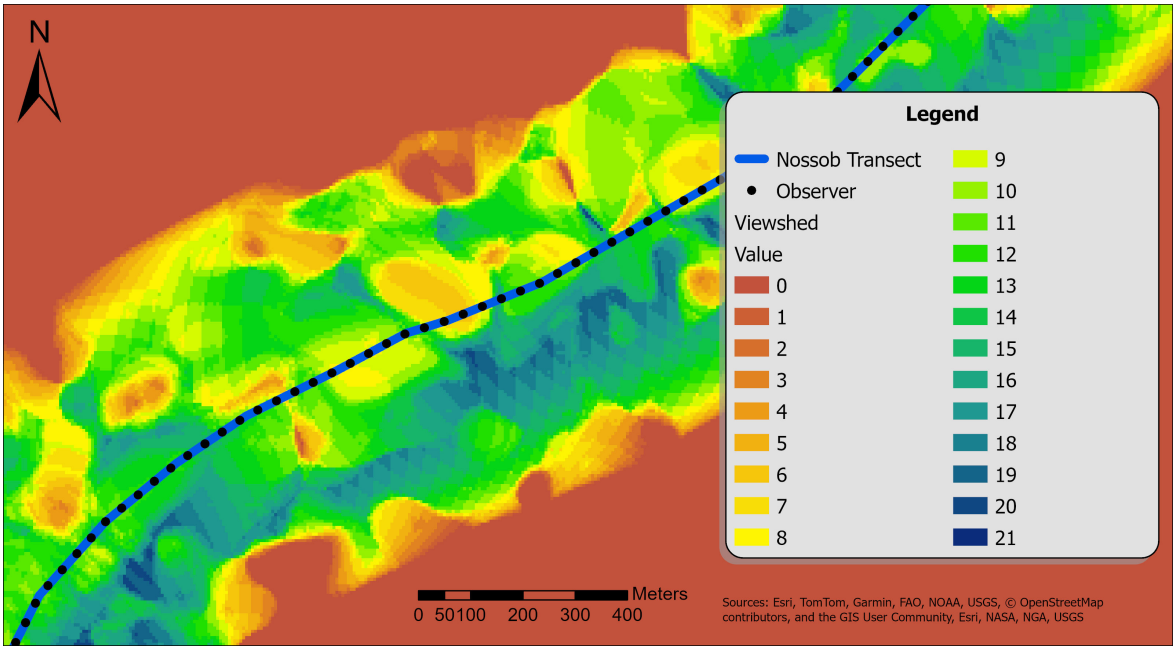
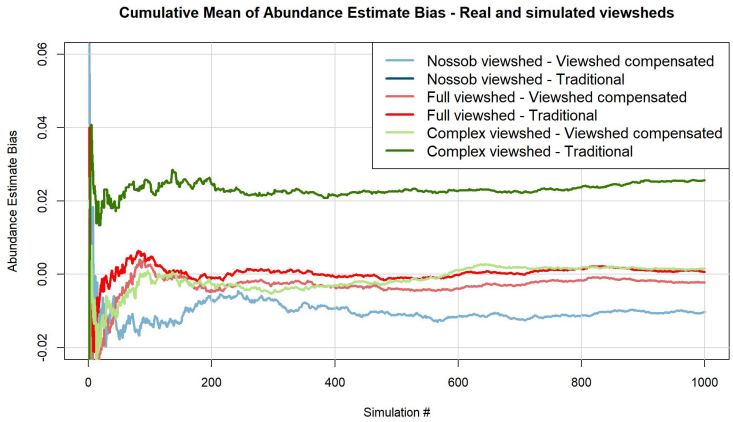


Figure 3.11: Part of the viewshed of the Nossob transect. The blue line is a part of the Nossob transect. Black dots are observers spaced 35 metres apart to calculate the viewshed. The viewshed value indicates how many regularly spaced observers were able to see raster cell. Maximum viewshed distance is 350 metres. Raster cell resolution is 5 metres

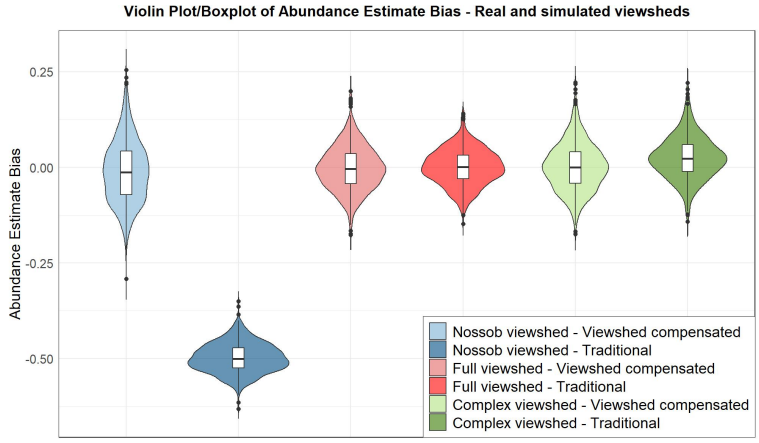
Viewshed compensated distance sampling without cluster sizes used is compared to traditional distance sampling, for the Nossob viewshed, the full viewshed, and the complex viewshed, in Figure 3.12 and Table 3.10.

Table 3.10: Comparison of real Nossob viewshed and simulated full and complex viewshed. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. No cluster sizes used. Number of simulations = 1.000, population size = 1.000. Input $\sigma = 180$.

Method	mean	median	sd	mean of last 10%
Nossob viewshed, viewshed compensated	-0.01023	-0.01293	0.08148	-0.01037
Nossob viewshed, traditional	-0.49816	-0.50081	0.03810	-0.49822
Full viewshed, viewshed compensated	-0.00229	-0.00421	0.057814	0.00208
Full viewshed, traditional	0.00062	0.00085	0.04699	0.00090
Complex viewshed, viewshed compensated	0.00159	-0.00008	0.06173	0.00142
Complex viewshed, traditional	0.02562	0.02324	0.05601	0.02542



(a) Cumulative mean of abundance estimate bias. Nossob viewshed traditional distance sampling outside plot extent.



(b) Violin plot and boxplot of abundance estimate bias.

Figure 3.12: Comparison of real Nossob viewshed and simulated full and complex viewshed. Methods applied: Viewshed compensated distance sampling and traditional distance sampling. No cluster sizes used. Number of simulations = 1.000, population size = 1.000. Input $\sigma = 180$.

For all methods in Figure 3.12 and Table 3.10, the cumulative mean of the abundance estimate bias has converged.

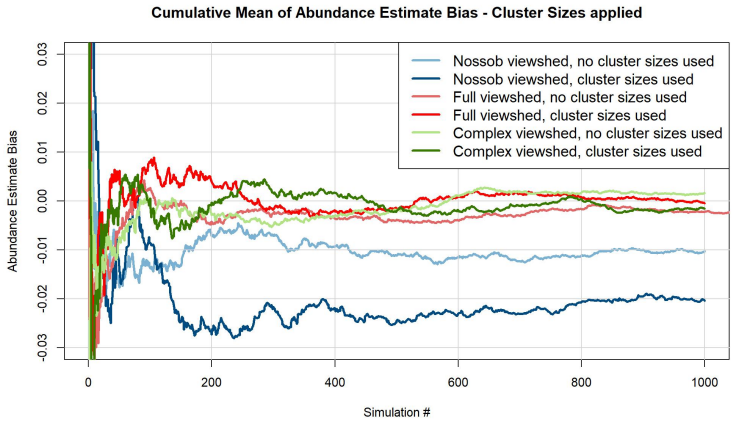
For each viewshed, viewshed compensated distance sampling performs better than traditional distance sampling. The mean abundance estimate bias is -0.49816 for traditional distance sampling, whereas the mean abundance estimate bias is -0.01023 for viewshed compensated distance sampling. This indicates a large increase in performance from traditional distance sampling to viewshed compensated distance sampling for the Nossob viewshed.

Given the means and standard deviations for the Nossob viewshed in Table 3.10 and the population size of 1.000, the abundance estimate with 95% confidence interval for viewshed compensated distance sampling is $\hat{N} = 990 \pm 158$ and the abundance estimate with 95% confidence interval for traditional distance sampling is $\hat{N} = 502 \pm 37$. The abundance estimate made using viewshed compensated distance sampling increased by $97.2\% \pm 31.7\%$, compared to the abundance estimate made using traditional distance sampling. For the complex viewshed the abundance estimate with 95% confidence interval for viewshed compensated distance sampling is $\hat{N} = 1002 \pm 121$.

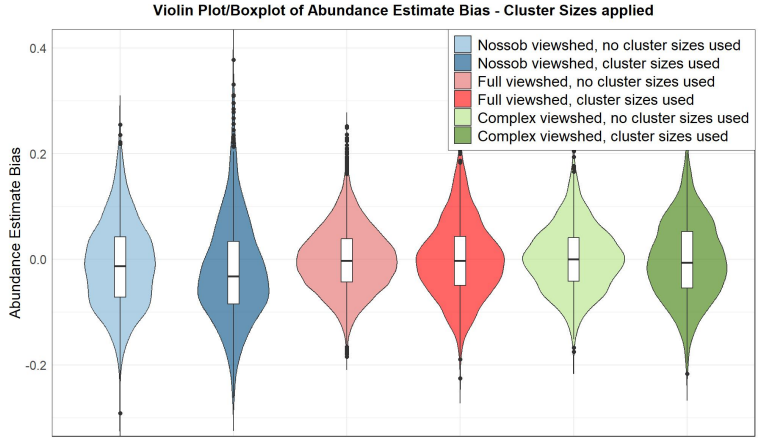
Viewshed compensated distance sampling with cluster sizes is compared to viewshed compensated distance sampling without cluster sizes, for the Nossob viewshed, the full viewshed, and the complex viewshed, in Figure 3.13 and Table 3.11.

Table 3.11: Comparison of Nossob viewshed, full viewshed and complex viewshed. With and without cluster sizes used. Method applied: Viewshed compensated distance sampling. Number of simulations = 1.000, population size = 1.000. Input $\sigma = 180$.

Method	mean	median	sd	mean of last 10%
Nossob viewshed, no cluster sizes used	-0.01023	-0.01293	0.08148	-0.01037
Nossob viewshed, cluster sizes used	-0.02042	-0.03221	0.09757	-0.01991
Full viewshed, no cluster sizes used	-0.00065	-0.00276	0.06028	-0.00085
Full viewshed, cluster sizes used	-0.00047	-0.00278	0.07223	0.00009
Complex viewshed, no cluster sizes used	0.00159	-0.00008	0.06173	0.00142
Complex viewshed, cluster sizes used	-0.00159	-0.00626	0.07366	-0.00181



(a) Cumulative mean of abundance estimate bias.



(b) Violin plot and boxplot of abundance estimate bias.

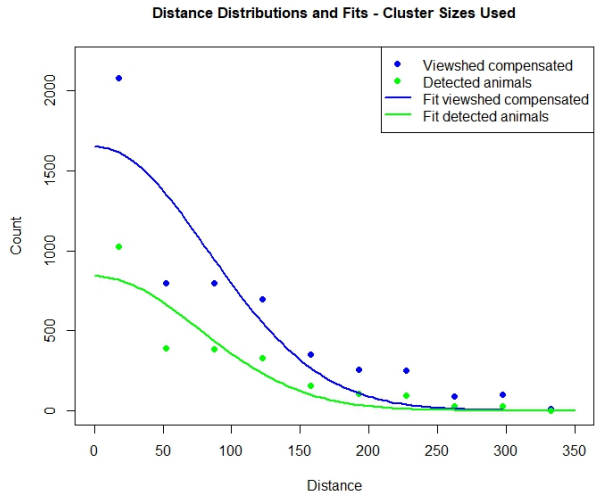
Figure 3.13: Comparison of Nossob viewshed, full viewshed and complex viewshed. With and without cluster sizes used. Method applied: Viewshed compensated distance sampling. Number of simulations = 1.000, population size = 1.000. Input $\sigma = 180$

For all methods in Figure 3.13 and Table 3.11, the cumulative mean of the abundance estimate bias has converged.

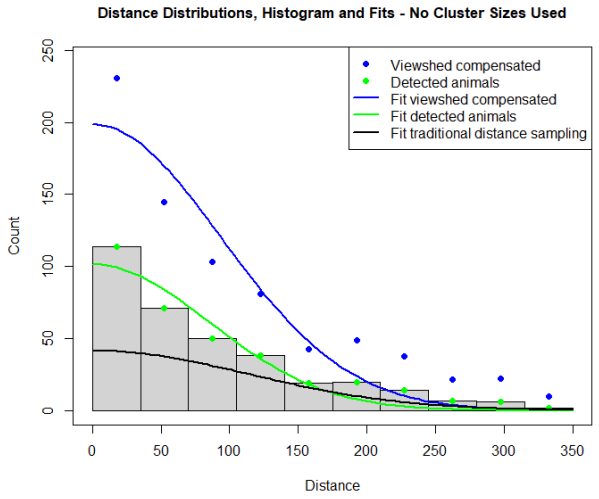
For the full viewshed and the complex viewshed, using cluster sizes has little impact on the performance. The standard deviation increases slightly. However for the Nossob viewshed the performance is lower when using cluster sizes.

3.3.3 Real input data

In Figure 3.14 the viewshed compensated distance distribution of the real animal population with the Nossob viewshed is shown, together with the distance distribution of detected animals. The distributions are shown with cluster sizes used and without cluster sizes used. In Table 3.12, the resulting detection probability estimates, abundance estimates and 95% prediction intervals can be found.



(a) Cluster sizes used. For viewshed compensated distance distribution (blue): $\sigma = 76.39$, $\tau = 839.79$. For detected animals distance distribution (green): $\sigma = 82.54$, $\tau = 1652.23$.



(b) No cluster sizes used. Histogram of distances of detected animals, and in black the fit. For viewshed compensated distance distribution (blue): $\sigma = 85.13$, $\tau = 101.48$. For detected animals distance distribution (green): $\sigma = 93.41$, $\tau = 198.79$. For traditional distance sampling (black), $\sigma = 114.50$, $\tau = 83.36$.

Figure 3.14: Results Nossob viewshed and real animal population. Viewshed compensated distance distribution and distance distribution of detected animals, with corresponding fit. Comparison between using cluster sizes and not using cluster sizes. In blue the viewshed compensated distance distribution and fit. In green, the distribution of detected animals and fit.

Table 3.12: Overview results Nossob viewshed and real animal population. Detection probability, abundance and lower 95% prediction interval (LPI) and upper 95% prediction interval (UPI). 95% prediction interval not available for detected animals results.

Method	Detection probability ($\hat{P}_a$)	Abundance ($\hat{N}$)	LPI	UPI
Viewshed compensated, cluster sizes used	0.2958	18367	14853	21881
Baseline, cluster sizes used	0.2735	9293	-	-
Viewshed compensated, no cluster sizes used	0.3344	2224	1869	2579
Baseline, no cluster sizes used	0.3048	1119	-	-
Traditional, no cluster sizes used	0.4091	833	771	895

With cluster sizes used the abundance estimate made using viewshed compensated distance sampling increased by 97.6%, compared to the abundance estimate made using baseline distance sampling. Without cluster sizes used the abundance estimate made using viewshed compensated distance sampling increased by 98.7%, compared to the abundance estimate made using baseline distance sampling. The similarity in this percentage indicates that using cluster sizes increases the absolute value of the abundance estimate, while keeping the same relative increase in abundance estimate.

The abundance estimate made using viewshed compensated distance sampling increased by $167.0\% \pm 42.7\%$, compared to the abundance estimate made using traditional distance sampling, when not using clusters. Only a small part of the 95% prediction interval overlaps with the 95% prediction interval of the percentage obtained for mixed input data, $97.2\% \pm 31.7\%$. This indicates that results from viewshed compensated distance sampling applied to mixed input data cannot directly be translated to results from viewshed compensated distance sampling applied to real input data.

The count in the first distance bin is very large compared to the counts in other distance bins. This indicates that relatively many animals are found within 35 metres from the transect. Due to the high count in the first distance bin, the detection function does not accurately fit the distribution.

4 Discussion

I developed viewshed compensated distance sampling, a method that explicitly uses the viewshed to compensate for the distance distribution of detected animals. My method improved the degree to which distance sampling satisfies assumptions 3 and 4 when the terrain is heterogeneous.

4.1 Methods discussion

Both Bayesian distance sampling and viewshed compensated distance sampling construct a single PDF, or mean viewshed per distance bin for viewshed compensated distance sampling. Furthermore, both methods use the perpendicular distance from the animal to the transect. Therefore, viewshed compensated distance sampling does not address the first two shortcomings of Bayesian distance sampling mentioned in the introduction. Viewshed compensated distance sampling addresses the third shortcoming, by modifying the interpretation of the viewshed. Whereas for Bayesian distance sampling the viewshed is interpreted as animals not being present in the blocked viewshed areas, for viewshed compensated distance sampling the viewshed indicates the chance of an animal being present in the area. This allowed the abundance estimate to cover the full area, rather than only the sampled area.

Due to the Bayesian statistics nature of Bayesian distance sampling, a confidence interval can be extracted for a single abundance estimate. For viewshed compensated distance sampling, simulations must be run with a simulated animal population to get a confidence interval. Therefore, for an abundance estimate made with a real viewshed and a real animal population, Bayesian distance sampling is able to define a confidence interval, while viewshed compensated distance sampling can only define a confidence interval based on simulations run with the real viewshed and simulated animal populations.

Viewshed compensated distance sampling directly incorporates cluster sizes by counting all animals in a distance bin rather than counting individual animal observations. In Bayesian distance sampling cluster sizes can be incorporated by estimating the mean cluster size and multiplying the mean cluster size posterior by the posterior distribution of abundance to derive a posterior distribution for the total abundance of animals (Delaney et al., 2025). Another method to incorporate cluster sizes into distance sampling is covariate distance sampling, where the detection function is modelled as a function of both distance and one or more additional covariates. The detection probability at a given distance and set of covariate values is denoted as $g(x, z)$, with z being the covariates involved. Examples of covariates are habitat, time of day, animal sex or cluster size (Buckland et al., 2004).

Viewshed compensated distance sampling has some other shortcomings. Bundling animals into distance bins is an irreversible operation that loses information, as individual distances can no longer be extracted. However, this operation is required to be able to compensate the distance distribution with the viewshed parameter distribution. Furthermore, bundling animals into distance bins introduces the variable of the distance bin width. Choosing a smaller distance bin width allows for more detail in the distance distribution. However, it also requires more data points to produce a smooth distribution. A smooth distribution is required to accurately fit the distance distribution. With a low number of observation, the distribution becomes very irregular, with a few individual animal counts having a large impact on the distance distribution.

When using simulated input data, the abundance estimate converged across many simulations.

However, for real input data, only a single distance distribution was available. When this distance distribution is too irregular, no abundance estimate can be obtained. This issue arises predominantly for small population sizes, and with cluster sizes applied, since the variability due to randomness is higher.

The viewshed compensated distance sampling method can be generalised, due to the way viewsheds are incorporated. Any viewshed raster with values indicating the likelihood that an animal is detected can be used as a viewshed. It is important that the assumptions on which the method is based are satisfied.

The results of viewshed compensated distance sampling on real input data cannot be generalised. The results are very specific to the Nossob riverbed and the game counts.

It was shown that viewshed incorporated distance sampling results in abundance estimates which are larger than abundance estimate by traditional distance sampling. These abundance estimates by viewshed incorporated distance sampling were more accurate than the abundance estimates by traditional distance sampling. However, viewshed incorporated distance sampling cannot directly replace distance sampling, especially when monitoring trends in wildlife abundance over time. Applying viewshed incorporated distance sampling does not indicate that wildlife abundance has increased, rather, previous abundance estimates were too low. Furthermore, it is not possible to directly apply viewshed incorporated distance sampling to abundance estimates made in the past, since often these abundance estimates were made using datasets without a distance attribute.

The fitting viewshed parameter distance sampling was developed as explained in appendix A.3. The goal of this method was to incorporate the viewshed parameter into the detection function. However, after developing this method, I realised that it was not possible to derive an abundance estimate from fitting both the distance distribution and the viewshed parameter distribution. In traditional distance sampling going from the distance distribution to an abundance estimate is possible because the distribution with an incomplete viewshed can be related to a complete viewshed through the detection probability. This cannot be translated to the viewshed parameter distribution, as a complete viewshed results in all viewshed parameters being 1, which does not allow any meaningful fitting. Therefore no baseline could be defined against which to compare the viewshed parameter distribution of an incomplete viewshed.

4.2 Input data discussion

A simple method was used for manually constructing the complex viewshed. Alternatively, a DEM could have been manually created and used as input for the method calculating the viewshed as described in section 2.6.2. This would allow for a better comparison between the performance of viewshed incorporated distance sampling with a simulated viewshed and the performance of viewshed incorporated distance sampling with a real viewshed.

For the simulated viewshed, the calculation of the mean viewshed per distance bin is straightforward, as long as a multiple of the distance bin width exactly equals the resolution of the raster cells. For a real viewshed, where the transect is not straight, the calculation of the mean viewshed per distance bin is more complicated, due to raster cells not being arranged systematically with respect to the transect.

In some parts of the Nossob viewshed shown in Figure 3.11, circular artefacts are seen. These are caused when multiple observers next to each other are all consistently able to see the area. This happens for large open areas, where the maximum viewing distance for observers is reached. Similar

to the smoothing applied by Brughmans et al. (2018) to address highly visible effects of contour artefacts, smoothing can be applied to the viewshed to resolve the artefacts. The effect of the circular artefacts can also be minimised by decreasing the distance between observer locations. The minimum distance between observer locations is 5 metres, determined by the resolution of the DEM. The disadvantage of decreasing the distance between observer locations is the increase in computation time for the viewshed.

The uniform distribution used to model cluster sizes does not accurately represent the true distribution of cluster sizes. By analysing the distribution of cluster sizes from the real animal population, a model can be defined to more accurately model cluster sizes for the simulated animal population. It was shown that the proportional increase in abundance estimate from baseline distance sampling to viewshed compensated distance sampling was similar regardless of whether cluster sizes were used. This indicates that, although the cluster sizes model may not accurately represent the true distribution of cluster sizes, it functions well for the real input data used in this thesis.

For my method applied to the Nossob viewshed and the real game counts, assumption 3 was not satisfied. Animals were not distributed independently of the transect, since animals were drawn to the riverbeds and boreholes. If the transects were located randomly, assumption 3 would still be satisfied (Burnham et al., 1980). This was not the case, with transects being driven through the riverbed, to allow for direct comparison of riverbed abundance and dune abundance.

Animals or animal clusters may have moved between the first moment of observation and the moment of measurement. In this case, the observer measured the distance to the location where the animal (cluster) was first observed. Inaccuracies were expected with this action. Additionally, for animal clusters, the observer manually estimated the cluster centre. This cluster centre was defined to minimise the weighted mean of distance to the cluster centre. However, for large animal clusters it was hard to accurately estimate the cluster centre. When animal clusters were larger than the 5 by 5 metres cell size, the viewshed parameter assigned to the animal cluster did not reflect the viewshed parameter for individual animals. All of these shortcomings violated assumption 5.

The real animal population obtained from the game counts had some shortcomings that violated the assumptions. Bearing was measured using a compass. Different compasses were used, such as a mechanical compass, an on-device compass and the compass in the app used to record game counts. These compasses produced readings that differed by up to 50 degrees. Using these bearings to calculate the perpendicular distance to transect violated assumption 6.

For the abundance estimate in the study area, game counts along the Nossob riverbed transect were used. These counts came from different days, times of day and observers. Additionally, some parts of transects driven along the Nossob riverbed transect were repeated. These repeated transects were included in the data, which may have skewed the results. These reasons meant that the abundance estimate could not yet be taken as the true abundance yet. First, more detailed analysis of the game counts data would have to be performed. The current abundance estimate gives a rough indication of the true abundance. Most importantly it shows the difference between incorporating viewsheds and not incorporating viewsheds.

5 Conclusion

I successfully developed a method that incorporates viewsheds into distance sampling. My method improves the performance in abundance estimation over traditional distance sampling. The performance of my method was compared to the performance of other methods according to the following research objectives.

Research objective 1 - simulated input data:
With the complex viewshed, without cluster sizes used, the abundance estimate improved from $\hat{N} = 9600 \pm 631$ for traditional distance sampling to $\hat{N} = 9974 \pm 741$ for my method.

With the irregular viewshed, without cluster sizes used, the abundance estimate was similar for my method ($\hat{N} = 10022 \pm 2321$) and Bayesian distance sampling ($\hat{N} = 9974 \pm 2084$). My method improves on Bayesian distance sampling by estimating abundance for the full area of interest rather, than the sampled area only.

Research objective 2 - mixed input data:
With the Nossob viewshed, without cluster sizes used, the abundance estimate improved from $\hat{N} = 502 \pm 37$ for traditional distance sampling to $\hat{N} = 990 \pm 158$ for my method.

With my method, without cluster sizes used, the abundance estimate for the Nossob viewshed ($\hat{N} = 990 \pm 158$) was worse than the abundance estimate for the simulated complex viewshed ($\hat{N} = 1002 \pm 121$).

Research objective 3 - real input data:
With the Nossob viewshed and the real animal population, with cluster sizes, the abundance estimate by my method was 97.6% larger than the abundance estimate by baseline distance sampling.

With the Nossob viewshed and the real animal population, without cluster sizes used, the abundance estimate by my method was $167.0\% \pm 42.7\%$ larger than the abundance estimate by traditional distance sampling. With a simulated animal population the abundance estimate by my method is $97.2\% \pm 31.7\%$ larger than the abundance estimate by traditional distance sampling. The increase in abundance estimate differed between the real and simulated animal populations, indicating that results from mixed input data cannot by directly extrapolated to real input data.

6 Recommendations

The game counts data includes attributes with the age and sex categories of animals and the animal species. Estimating abundance by species, age, and sex category would provide insight into trends across these groups.

To accurately evaluate the performance of incorporating cluster sizes into viewshed compensated distance sampling, cluster sizes must also be incorporated into distance sampling as a baseline for comparison. Established methods are available for this purpose. These methods should be implemented to allow viewshed incorporated distance sampling to be compared against them. Additionally, the cluster sizes model should be refined to more accurately reflect the true distribution of cluster sizes in real animal populations.

Use of generative Artificial Intelligence

During the preparation of this work the author used Claude Sonnet 4.6 (accessible via: www.claude.ai) and ChatGPT GPT-5.3 (accessible via: www.chatgpt.com) in order to write R code, to set up the Overleaf document, to assist with ArcGIS workflows, to assist with brainstorming, to obtain feedback on the general flow of the text, and to obtain feedback on spelling and grammar. After using these services, the author reviewed and edited the content as needed. The author takes full responsibility for the content of the publication. Some examples of the user-prompts and model outputs can be found in appendix A.5.

During the preparation of this work the author used FeedbackFruits through Brightspace to receive feedback on the report in areas such as grammar, spelling, punctuation and citation style and. After using these services, the author reviewed and edited the content as needed. The author takes full responsibility for the content of the publication.

References

ArcGIS (2026). Geodesic Viewshed (Spatial Analyst)—ArcGIS Pro | Documentation.

Batter, T. J., Landers, R. H., Denryter, K., and Bush, J. P. (2022). Use of aerial distance sampling to estimate abundance of tule elk across a gradient of canopy cover and comparison to a concurrent fecal DNA spatial capture-recapture survey.

Brughmans, T., van Garderen, M., and Gillings, M. (2018). Introducing visual neighbourhood configurations for total viewsheds. *Journal of Archaeological Science*, 96:14–25.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2004). *Advanced Distance Sampling: Estimating abundance of biological populations*. OUP Oxford. Google-Books-ID: hBRREAAAQBAJ.

Buckland, S. T., Rexstad, E. A., Marques, T. A., and Oedekoven, C. S. (2015). *Distance Sampling: Methods and Applications*. Methods in Statistical Ecology. Springer International Publishing, Cham.

Burnham, K. P., Anderson, D. R., and Laake, J. L. (1980). Estimation of Density from Line Transect Sampling of Biological Populations. *Wildlife Monographs*, (72):3–202.

Callaghan, C. T., Santini, L., Spake, R., and Bowler, D. E. (2024). Population abundance estimates in conservation and biodiversity research. *Trends in Ecology & Evolution*, 39(6):515–523.

Carswell, B. M., Avgar, T., Street, G. M., Boyle, S. P., and Vander Wal, E. (2025). One size does not fit all: A novel approach for determining the Realised Viewshed Size for remote camera traps. *Methods in Ecology and Evolution*, 16(4):786–800. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.70008.

Craigie, I. D., Baillie, J. E. M., Balmford, A., Carbone, C., Collen, B., Green, R. E., and Hutton, J. M. (2010). Large mammal population declines in Africa's protected areas. *Biological Conservation*, 143(9):2221–2228.

CyberTracker (2026). CyberTracker.

Delaney, D. M., Harms, T. M., Harris, J. P., Kaminski, D. J., Elliott, J. R., and Dinsmore, S. J. (2025). Methods to account for incomplete viewsheds in distance sampling. *Methods in Ecology and Evolution*, 16(6):1201–1214. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.70038.

Dodge, Y. (2008). Bernoulli Distribution. In *The Concise Encyclopedia of Statistics*, pages 36–37. Springer, New York, NY.

Dubey, A. (2023). Species abundance | Definition, Conservation, & Facts | Britannica.

EPSG (2026). Hartebeesthoek94 / Lo29 - EPSG:2053.

Fisher, P. (1991). First Experiments in Viewshed Uncertainty: The Accuracy of the Viewshed Area. *Photogrammetric Engineering & Remote Sensing*, 57.

Gates, C. E., Marshall, W. H., and Olson, D. P. (1968). Line Transect Method of Estimating Grouse Population Densities. *Biometrics*, 24(1):135–145.

IPBES, Díaz, S., Settele, J., Brondízio, E. S., Ngo, H. T., Guèze, M., Agard, J., Arneth, A., Balvanera, P., Brauman, K. A., Butchart, S. H. M., Chan, K. M. A., Garibaldi, L. A., Liu, K. I. J., Subramanian, S. M., Midgley, G. F., Miloslavich, P., Molnár, Z., Obura, D., Pfaff, A., Polasky, S., Purvis, A., Razzaque, J., Reyers, B., Chowdhury, R. R., Shin, Y. J., Visseren-Hamakers, I. J., Willis, K. J., and Zayas, C. N. (2019). Summary for policymakers of the global assessment report on biodiversity and ecosystem services. Technical report, Zenodo. Version Number: summary for policy makers.

Kaminski, D. J., Harms, T. M., and Coffey, J. M. (2019). Using spotlight observations to predict resource selection and abundance for white-tailed deer. *The Journal of Wildlife Management*, 83(7):1565–1580.

Kansanen, K., Packalen, P., Maltamo, M., and Mehtätalo, L. (2021). Horvitz-Thompson–Like Estimation with Distance-Based Detection Probabilities for Circular Plot Sampling of Forests. *Biometrics*, 77(2):715–728.

Kellner, K. and Meredith, M. (2026). jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses.

Koeck, M. M., Moeller, A. K., Elmore, R. D., Fairbanks, W. S., and Chitwood, M. C. (2025). Accounting for Viewshed Area and Animal Availability When Estimating Density and Recruitment of Unmarked White-Tailed Deer. *Ecology and Evolution*, 15(3):e71015.

Maichak, E. J. and Schuler, K. L. (2004). Applicability of Viewshed Analysis to Wildlife Population Estimation. *The American Midland Naturalist*, 152(2):277–285.

Morgan, D. G., Gibbens, J. R., Conway, E. T., Hepworth, G., and Clough, J. (2024). Wildlife density estimation by distance sampling: A novel technique with movement compensation. *PLOS ONE*, 19(10):e0310020.

Moswete, N. N., Thapa, B., and Child, B. (2012). Attitudes and opinions of local and national public sector stakeholders towards Kgalagadi Transfrontier Park, Botswana. *International Journal of Sustainable Development & World Ecology*, 19(1):67–80. _eprint: https://doi.org/10.1080/13504509.2011.592551.

Norton-Griffiths, M. (1978). Counting Animals. *Counting animals: Revised second edition. Handbook No. 1.*

Okorondu, J., Nwaogu, C. O. U., and Umar, N. A. (2022). Anthropogenic Activities as Primary Drivers of Environmental Pollution and Loss of Biodiversity: A Review. *International Journal of Trend in Scientific Research and Development*, Volume-6(Issue-4).

Pacey, P. (2026). Citizen Science and Game Counts.

Plummer, M. (2003). JAGS: A Program for Analysis of Bayesian Graphical Models using Gibbs Sampling. *3rd International Workshop on Distributed Statistical Computing (DSC 2003); Vienna, Austria*, 124.

Prakash, S. and Verma, A. K. (2022). Anthropogenic Activities and Biodiversity Threats.

Robert, C. and Casella, G. (2011). A Short History of Markov Chain Monte Carlo: Subjective Recollections from Incomplete Data. *Statistical Science*, 26(1):102–115.

Royle, J. A. (2004). N-Mixture Models for Estimating Population Size from Spatially Replicated Counts. *Biometrics*, 60(1):108–115.

Teunissen, P. J. G. (1990). Nonlinear least squares. *manuscripta geodaetica*, 15(3):137–150.

Thomas, L., Buckland, S. T., Rexstad, E. A., Laake, J. L., Strindberg, S., Hedley, S. L., Bishop, J. R., Marques, T. A., and Burnham, K. P. (2010). Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology*, 47(1):5–14. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2664.2009.01737.x.

White, G. C. (1982). *Capture-recapture and Removal Methods for Sampling Closed Populations*. Los Alamos National Laboratory. Google-Books-ID: dHPwAAAAMAAJ.

Williamson, D. and Williamson, J. (1984). Botswana's fences and the depletion of Kalahari wildlife. *Oryx*, 18(4):218–222.

A Appendix

A.1 Variables and parameters

Table A.1: Definition of variables and parameters used in thesis

Parameter	Unit	Explanation
N	-	Animal abundance
A	m^2	Area of interest
n	-	Number of counted animals
w	m	Half-width of area of interest
L	m	Length of the transect
x	m	Distance from transect to animal (cluster)
v	-	Viewshed parameter
s	-	Cluster size
P_a	-	Detection probability
σ	-	Scale parameter in the half-normal function controlling steepness
τ	-	Parameter in viewshed compensated distance sampling controlling value of half-normal function at $x = 0$
ψ	-	Parameter in Bayesian distance sampling representing the probability of an individual corresponding to a real animal
α	-	Parameter in fitting viewshed parameter distance sampling controlling value of fit at $v = 0$
β	-	Parameter in fitting viewshed parameter distance sampling controlling steepness of fit
ρ	-	Parameter controlling the slope of the linear part of the modified half-normal function
$g(x)$	-	Detection function
$f(x)$	-	Probability density function of distances of detected animals
$\pi(x)$	-	Probability density function of distances of detected and undetected animals

A.2 CyberTracker

CyberTracker is a platform which can be used to build mobile applications to record, visualise and analyse observations in nature (CyberTracker, 2026). On the website a project can be developed. This project is easily opened on the mobile app. An application was developed to keep track of the game counts. A user guide has been developed to guide the user through the app.

The user (and their team) drove their vehicle and observed the environment until an animal cluster was detected (capture event). When an animal cluster was detected the vehicle stopped and the user measured and recorded the following three items: (1) distance to the centre of the animal cluster, (2) bearing to the centre of the animal cluster and (3) number of animals in the capture event per age/sex category. The distance was measured using a laser rangefinder and the bearing was measured

using a compass. In the background the GPS location was recorded. Only herbivores with a body mass exceeding five kilograms were recorded.

A.3 Fitting viewshed parameter distance sampling

As outlined in section 2.1, the optimal σ is estimated based on the sampled distance data using maximum likelihood fitting. To incorporate viewsheds, the detection function is modified to not only include distance (x), but also include the viewshed parameter (v). The viewshed parameter is an animal specific value, describing the likelihood that an observer was able to see the animal from the transect, with $v \in [0, 1]$. The modified detection function to fit both distance and viewshed parameter is:

$$g(x, v) = \rho v e^{-\frac{x^2}{2\sigma^2}} \tag{A.1}$$

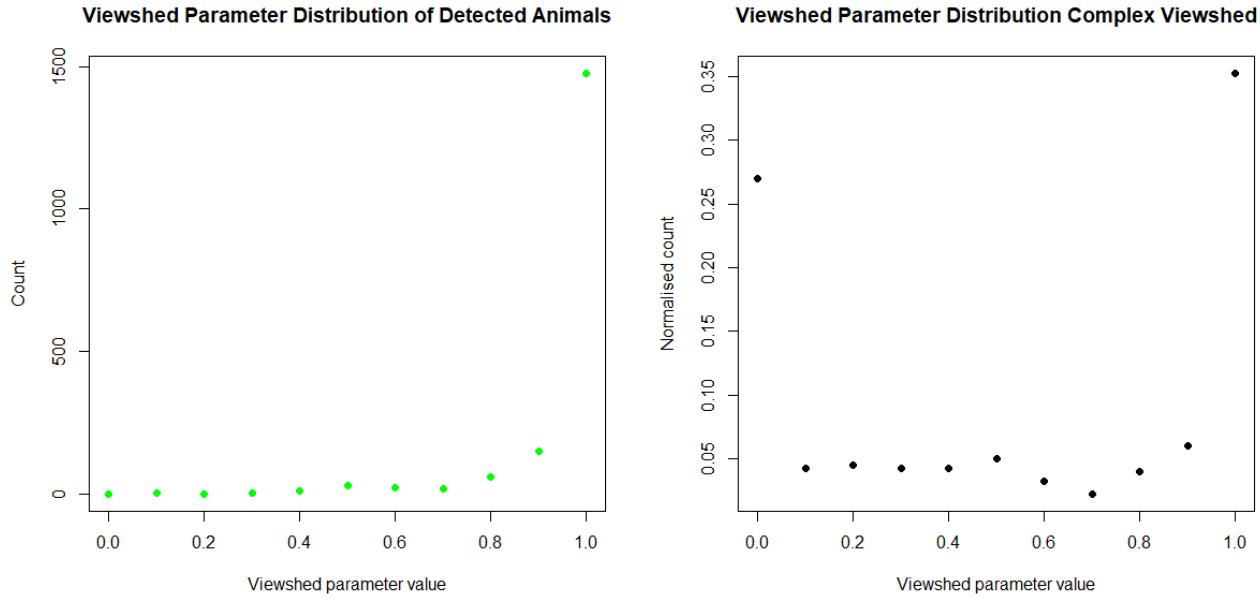
This function not only includes σ, but also ρ. ρ controls the slope of the linear relation between the output and the viewshed parameter. Since v and x are not directly linked in equation A.1, the function could be separated to handle v and x separately. For each animal the distance and the viewshed parameter are extracted and counted per viewshed parameter bin. An example of the distribution of viewshed parameters is shown in Figure A.1a. The distribution of viewshed parameters has to be corrected for the distribution of viewshed parameters available. If this is not done, viewshed parameters of detected animals are overrepresented with viewshed parameters which are found more commonly. An example of the distribution of available viewshed parameters, summing to 1, is shown in Figure A.1b.

Each viewshed parameter count of detected animals (Figure A.1a) is divided by the normalised viewshed parameter count of the viewshed raster (Figure A.1b). This effectively increases the count of animal detections with a low viewshed parameter.

In case the viewshed parameter distribution is not linear, a exponential relation can be explored:

$$g(x, v) = \alpha e^{\beta v} e^{-\frac{x^2}{2\sigma^2}} \tag{A.2}$$

The parameters of the new detection function can be used to calculate the abundance estimate. However, as explained in the discussion this is not possible.



(a) Viewshed parameter distribution of detected animals.

(b) Viewshed parameter distribution of all viewshed parameters.

Figure A.1: Example of distributions of viewshed parameters. Complex viewshed used (Figure 2.4e).

In Figure A.2 the results for the fitting viewshed parameter distance sampling method are shown.

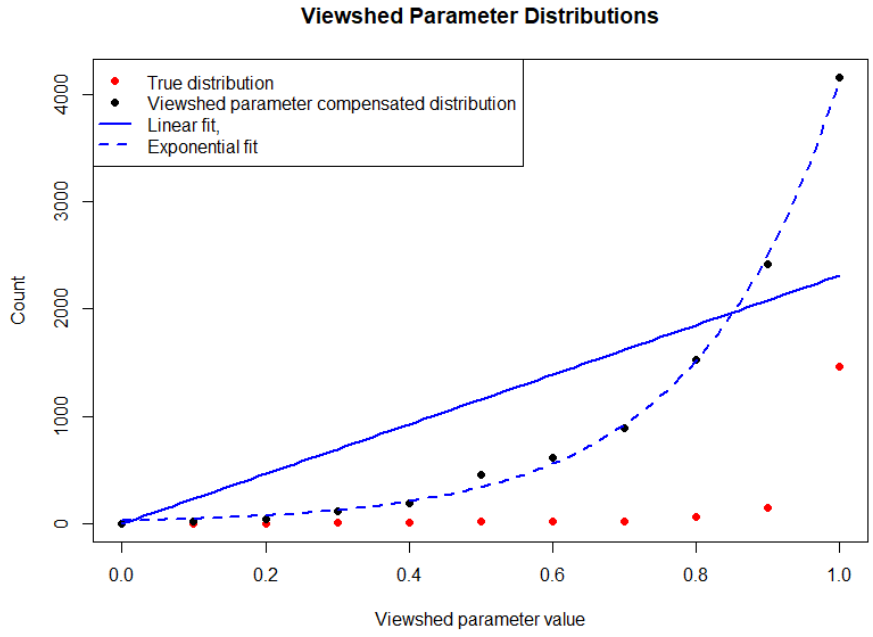


Figure A.2: Viewshed parameter distributions. In red, the viewshed parameter distribution. In black, the viewshed parameter distribution compensated for the available viewshed parameters. Solid blue line, linear fit with $\tau = 2310$. Dashed blue line, exponential fit with $\alpha = 27.92$, $\beta = 4.99$. Input $\sigma = 180$.

The exponential equation (equation A.2) does not go through the origin, although at $v = 0$ no

animals can be detected. Therefore the exponential detection function cannot be a perfect fit.

A.4 Research data

The code and data to reproduce the results of this thesis are available on `https://github.com/ThijsVons/Thesis-Viewsheds`. This includes the simulated viewsheds which were used in this thesis. It does not include the data of the real Nossob viewshed and the game counts.

Below is an overview of the research data which was submitted.

- README.txt
- Thesis report (pdf)
- Midterm presentation (pdf)
- Final presentation (pdf)
- Datasets (.csv)
- Figures (.png/.jpeg)
- Scripts (.R/.py)
- ArcGIS toolboxes (.atbx)

A.5 AI prompts

Claude, R code

Author: I have a dataframe of 270 rows and 6 columns. It has a row of names and a column of headers. I want a function in R. This function should split up the data in blocks of 27 rows (and 6 columns). It should take the average over all 10 blocks.

Claude: Here's an R function that will split your dataframe into blocks of 27 rows and calculate the average across all blocks:

```r
calculate_block_averages <- function(df) {
  # Remove the header row if it's included in the data
  # Assuming first row contains names that should be excluded
  data_only <- df[-1, ]

  # Convert to numeric matrix (in case columns are character type)
  data_matrix <- as.matrix(data_only)
  mode(data_matrix) <- "numeric"

  # Define block size
  block_size <- 27
  n_blocks <- 10

  # Initialize matrix to store results
  block_means <- matrix(0, nrow = block_size, ncol = ncol(data_matrix))

  # Calculate mean for each position across all blocks
  for (i in 1:n_blocks) {
    start_row <- (i - 1) * block_size + 1
    end_row <- i * block_size
    block_means <- block_means + data_matrix[start_row:end_row, ]
  }

  # Divide by number of blocks to get average
  block_means <- block_means / n_blocks

  # Convert back to dataframe with original column names
  result <- as.data.frame(block_means)
  colnames(result) <- colnames(df)

  return(result)
}
```

Figure A.3: Claude prompt response R code.

This function:

- Removes the first row (assuming it contains names)
- Converts the data to numeric format
- Splits the 270 rows into 10 blocks of 27 rows each
- Calculates the element-wise average across all 10 blocks
- Returns a dataframe with 27 rows and 6 columns containing the averaged values

If your header row is already properly set as column names (not a data row), you can remove the data_only <- df[-1,] line and just use df directly.

Author: The first row does not have to be removed. When loading the data I used header=TRUE and row.names=1. So the dataframe is just 270x6.

Claude: Perfect! Here's the simplified function for your 270x6 dataframe:

```r
calculate_block_averages <- function(df) {
  # Convert to numeric matrix
  data_matrix <- as.matrix(df)
  mode(data_matrix) <- "numeric"

  # Define block size
  block_size <- 27
  n_blocks <- 10

  # Initialize matrix to store results
  block_means <- matrix(0, nrow = block_size, ncol = ncol(data_matrix))

  # Calculate mean for each position across all blocks
  for (i in 1:n_blocks) {
    start_row <- (i - 1) * block_size + 1
    end_row <- i * block_size
    block_means <- block_means + data_matrix[start_row:end_row, ]
  }

  # Divide by number of blocks to get average
  block_means <- block_means / n_blocks

  # Convert back to dataframe with original column names
  result <- as.data.frame(block_means)
  colnames(result) <- colnames(df)

  return(result)
}
```

Figure A.4: Claude prompt response R code.

Usage:

```r
# Load your data
my_data <- read.csv("yourfile.csv", header = TRUE, row.names = 1)

# Calculate block averages
averaged_data <- calculate_block_averages(my_data)
```

Figure A.5: Claude prompt response R code.

The function returns a 27x6 dataframe with the averaged values across all 10 blocks, preserving your original column names.

ChatGPT, ArcGIS development

Author: I have a polyline in ArcGIS. I want to delete a part of it, so I want to split the line. I assume I should use the 'Split' tool. How do I use that tool?

ChatGPT: You're on the right track — yes, you split the polyline first, then delete the piece you don't want.
Below are the simple, practical steps, with notes for ArcGIS Pro and ArcMap (they're very similar).

Concept first (quick sanity check)
You cannot delete part of a single line directly. You must:

1. Split the line at the desired location(s)
2. Select and delete the unwanted segment line at the desired location(s)

ArcGIS Pro — Step by Step

1. Start an edit session
   (a) Go to the Edit tab
   (b) Click Modify
2. Choose the Split tool
   In the Modify Features pane:
   (a) Select Split
   (b) Icon looks like a line with a break
3. Select the line to split
   (a) Click the polyline you want to edit
   → it highlights
4. Split the line
   You have two common ways to split:
   Option A: Click split (most common)
   (a) Click once on the line where you want the split
   (b) Press Enter

   Option B: Draw a split line
   (a) Click Draw mode
   (b) Draw a line crossing the polyline
   (c) Double-click to finish

   The polyline is now two separate features
5. Delete the unwanted segment
   (a) Switch to the Select tool
   (b) Click the piece you don't want
   (c) Press Delete
6. Save edits
   Edit → Save Edits

ArcMap (classic) - Quick Version

1. Start editing: Editor → Start Editing
2. Select the line feature
3. Choose Editor → Split
4. Click where you want to split the line
5. Select unwanted segment → Delete
6. Edits

Pro tips (these save headaches)

- Turn on snapping if you need precision
- You can split a line multiple times before deleting
- Attributes are duplicated when splitting (important if you're editing networks)

If you want, tell me:

- ArcGIS Pro or ArcMap?
- Are you splitting at a point, between two points, or by distance?

I can tailor the exact workflow for your case.