**Exploratory data analysis**

The top left histogram of Figure 1 shows that the perpendicular distances are generally close to the transect line. While distances were measured accurately using a laser range finder, the angles were taken with a compass. Analysis of the angles measured in the bottom left histogram of Figure 1 shows that they have been rounded to the nearest 10 degrees, as well as angles under 20 being heaped to 0.

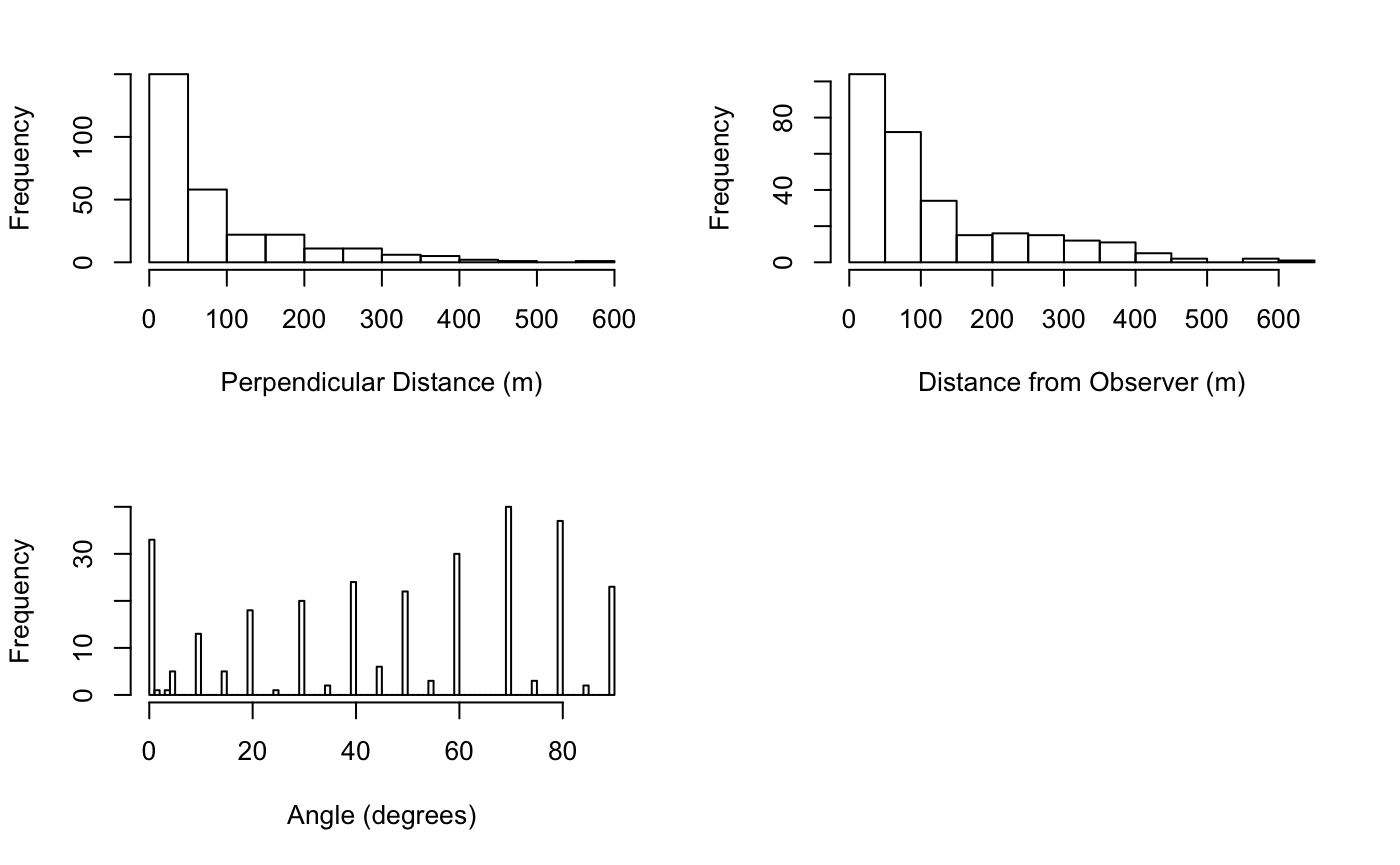


Figure 1 Analysis of Measurements taken during Survey

Buckland, 2001 (p267) suggests five ways of handling distance measurement error.

1. Record distances and angles more accurately
2. Use models for the detection function that always have a shoulder
3. Group the data before analysis
4. ‘Smear’ the data
5. Use radial distance models.

It is too late for 1, according to the text 5 leads to non-robust models. 2, 3 and 4 will be explored to find a good fitting model.

**‘Smearing’ the measured angles**

Buckland, 2001 (p269) says that although often criticized smearing is regularly used in cetacean surveys to handle errors in angle readings. Calculations used below are given by D. Kinzey, 2002 and based on the text in Buckland, 2001. The angles are adjusted within a given range using a uniform distribution

Sρ = | ρ + υ ∆ρ |

Where ρ = recorded angle, υ = a uniform random number between -0.5 and 0.5, and ∆ρ = the range of angles to be smeared over.

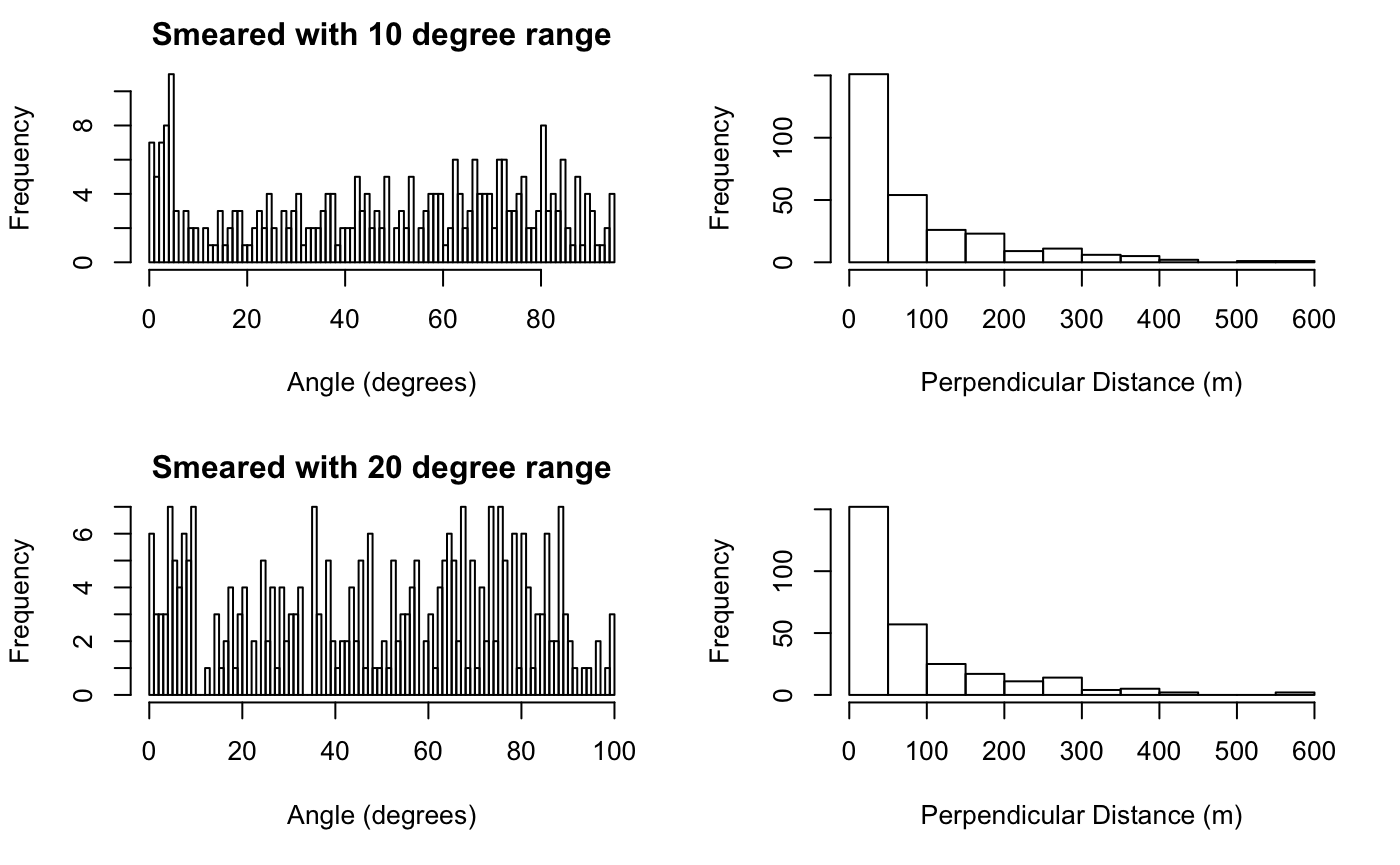


Figure Effects of Smearing the Measured Angle

Figure 2 shows that there is very little effect in the perpendicular distance when smearing the angle. Figure 3 shows that majority of the distances recorded at low angles were for animals very close to the observer, so the smeared angle has very little effect on the distance from the transect line.

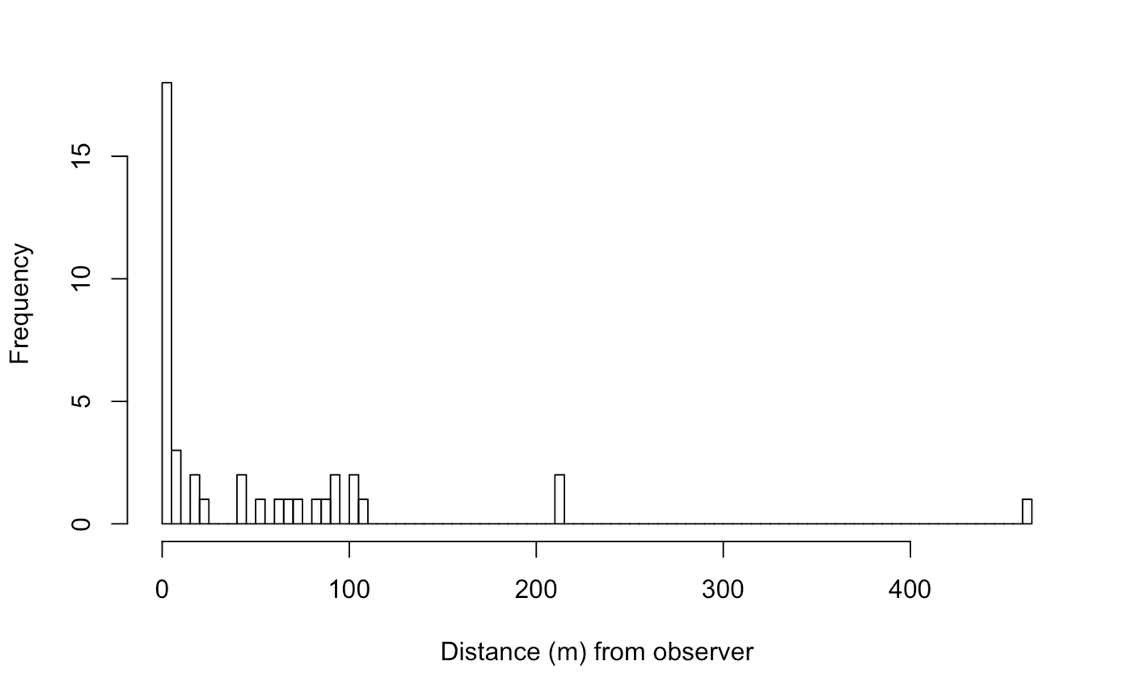


Figure Distance of Animals from Observer where measured angle is less than 10 degrees

**Grouping the data**

Buckland, 2001 (p109) says that grouping should be used where there are measurement errors and that appropriate grouping can lead to a better model fit and improved estimates of density. The text suggests cutpoints should be selected that avoid the rounding distances and when a lot of sightings fall on the transect (particularly common when using angles and distances) then a relatively wide first internal should be chosen. C Bibby, 1998 says that when grouping data to deal with heaping, the first interval chosen should be narrow and fall within the ‘shoulder’ and the others should increase with distance from the transect. Buckland, 2001 (p158) suggests that six to eight groups is a reasonable number or resolving heaping.

Choosing the cutpoints will be difficult as the heaping occurred with the angles measured rather than the distances.

**Shape Criterion**

Buckland, 2001 (p42) discusses the shape criterion and that a detection function should have a shoulder, which means that the derivative of the detection function on the transect is equal to 0, i.e. the observer detects all animals on and just off the transect. The text goes on to say that this is particularly important when heaping at zero is suspected. The exponential model will give unreliable probabilities at 0 and so will not meet this criterion.

**Cluster Size**

Larger cluster sizes are obviously more visible at distance than single animals. Thus there is a bias for only recording animals in large clusters at a far distance. Figure 4 shows that there is some cluster size bias in the hare data.

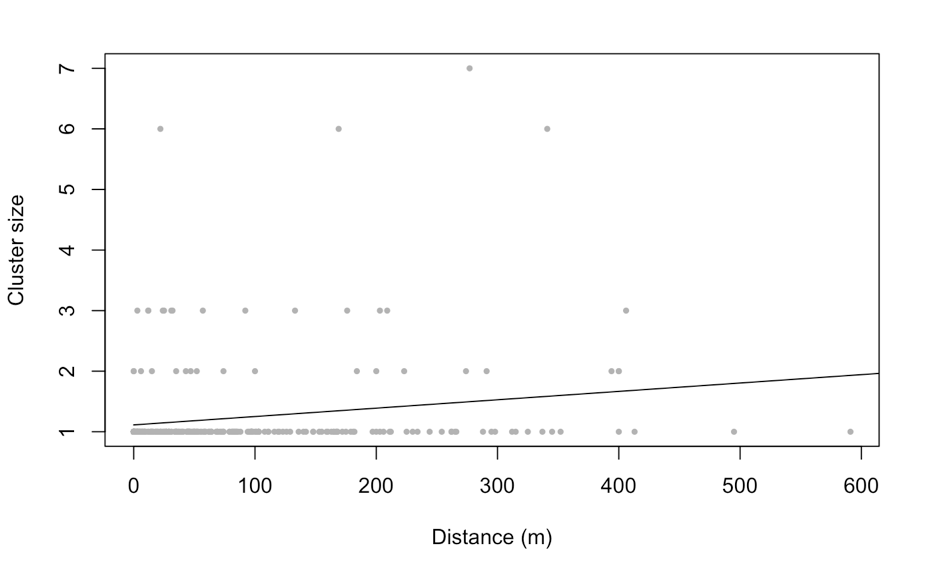


Figure Detection of Large Clusters at Distance

Approaches for dealing with cluster size bias include:

1. Truncate the data to ignore large distances where bias occurs.
2. Add the cluster size as a covariate to the distance when modelling.
3. Calculate the mean cluster size using a regression estimator.
4. Replace clusters with individuals. When the number of detected clusters is small (13% of detections were clusters in the hare data) then this approach may perform poorly.

In the hare data less clusters of 2 or 3 hares were seen from about 200m.

**Truncation**

Right truncating the data may lead to a little loss of precision, however, losing some outliers will help reduce the number of parameters in the detection function and so reduce bias (Buckland, 2001 (p107)).

Truncation approaches include:

1. Truncating the furthest 5% of sightings.
2. Fitting a preliminary model and truncating any sightings at a distance with less that 15% chance of detection.

**Model Selection**

The problem with working is that model fits are not always repeatable when comparing two different smears.

Many models that have a reasonable fit show that either far more than 100% of animals were seen at 0 or many less.

The model chosen has a good fit and shows that 100% of animals were seen at 0. The model has a shoulder; however, the shoulder sits very close to 0 and the first group does not sit within the shoulder. Cluster size bias is handled by having size as a covariate. The data is truncated at 400m which only removes 4 sightings.

The first group needs to be about 25m and you need a long tail off to create a good probability curve.

**Bibliography**

Buckland 2001

INTRODUCTION TO DISTANCE SAMPLING

D. Kinzey 2002

ACCURACY AND PRECISION OF PERPENDICULAR DISTANCE MEASUREMENTS IN SHIPBOARD LINE-TRANSECT SIGHTING SURVEYS

C Bibby 1998

EXPEDITION FIELD TECHNIQUES - BIRD SURVEYS