Review of Kangaroo Management Programme Surveys in New South Wales

Authors:	Borchers D, Burt L, Marshall L, Oedekoven C
Report Code:	CREEM-2023-04
Date:	October 2023

Please cite this report as: Borchers D, Burt L, Marshall L, Oedekoven C (2023) Review of Kangaroo Management Programme Surveys in New South Wales. Report number CREEM-2023-04. Provided to Department of Planning and Environment, NSW Government. Australia, October 2023 (unpublished).

Document Control

Please consider this document as uncontrolled copy when printed

Rev.	Date.	Reason for Issue.	Prepared by	Checked by
2	11/08/2023	Draft	LB, CSO, LHM, DLB	LB, DLB
3	07/10/2023	Final draft	LB, LHM, CSO, DLB	LB
4	24/10/2023	Final	LB, LHM, CSO, DLB	LB

Disclaimer

Executive summary

The New South Wales (NSW) Department of Planning, and Environment has researched, monitored, and managed the large kangaroo population under a Kangaroo Management Programme (KMP) since 1974 (NSW KMP, 2021). Surveying any species of wild animal to obtain reliable abundance estimates is a challenging and costly task, especially over large areas. Distance sampling is one of the most widely used methods for estimating animal abundance and is particularly effective at estimating population sizes over large areas. Aerial surveys have been undertaken annually by the KMP employing distance sampling methods; aircraft fly along pre-determined transects and observers search for kangaroos and record key information upon detection of each animal. Two separate surveys are carried out: a fixed-wing aircraft survey in the Western Plains and a helicopter survey in the Tablelands. The methods rely on certain assumptions being valid for reliable abundance estimates to be obtained. In situations where standard assumptions are not met, it may be possible to modify standard distance sampling methods to account for this.

The Centre of Ecology and Environmental Modelling (CREEM), University of St Andrews, was approached to conduct a review of the survey design, protocols and analysis methods used by the KMP. CREEM has a long history of developing distance sampling methods and tools (e.g., Thomas *et al*., 2010). This report contains the results of this review. The KMP reports and data were provided to CREEM and there were several meetings between the two groups to discuss various aspects of the surveys. The KMP engaged constructively with CREEM in the review process.

The report outlines the methods being used, to provide context for recommendations. It briefly describes the implementation of the methods that were used to obtain abundance estimates. Any aspects of concern that either CREEM or the KMP identified, are discussed and recommendations made. Technical details are contained in appendices.

Overall, we found the KMP to be very thorough in its approach, to have considered all relevant aspects of the surveys and to be willing to make improvements and consider new technologies to ensure that the abundance estimates are as reliable as possible. We detail below some areas in which we consider the KMP to be following best practice and to have taken steps to overcome the challenges presented by the whole survey process.

- In both Tablelands and Western Plains, the survey designs were randomised with the aim of sampling a representative area of the study region.
- Survey design in the Tablelands regions was stratified by expected animal density to improve the overall precision of the estimates.
- Distance for Windows software was used to assess the survey design coverage for the segmented grid design in the Tablelands.
- Consideration was given as to the orientation of the lines to avoid severe glare along one side of the transect; zigzag lines were selected in the Western Plains to overcome this issue.
- Mark recapture distance sampling (MRDS) techniques were employed in the Western Plains, where the probability of detecting animals on the transect may be less than one, to estimate the probability of missing animals on the transect.
- Substantial effort was dedicated to training observers, in particular, in the following areas: focussing effort close to the transect, defining clusters (in particular for Western Plains surveys where cluster definition is of higher importance), measuring to the centre of a cluster, rapid data entry.

- Accurate data collection was aided using dedicated software and data entry equipment in the Western Plains surveys and voice recording devices in the Tablelands. These aids minimise the time observers take their eyes off the transect, if at all, to record data.
- For Western Plain surveys, cluster sizes are recorded by observers via subitising instead of counting the total numbers to reduce errors.
- The aircraft (fixed-wing planes and helicopters) have distance-marked struts attached to aid the observers in correctly allocating detected animals/groups to distance bins.
- Careful consideration was given to the matching of the detections between observers on the Western Plains surveys, where MRDS methods were used, to determine which were re-sightings.
- A point independence mark recapture model was fitted to the Western Plains data as recommended for aerial kangaroo surveys by Fewster and Pople (2007).
- A variety of models (e.g., covariates considered in the detection function) to account for differences between detections.
- Industry-standard model selection and techniques to assess model assumptions and fit were considered during the analyses to select the most appropriate model.

The KMP has employed the most up-to-date and robust methods for population abundance estimation. In a few instances, logistical constraints have meant that some assumptions of the methods have not been met and we make recommendations to either better meet these assumptions or account for them in the abundance estimation. In addition, we advise on recently developed software, or potential software applications, for implementing these methods. Many suggestions for improvements can be implemented relatively easily but others may have substantial practical and/or cost implications. The main recommendations are listed below but note that not all points will relate to both the Western Plains and Tablelands surveys:

- We recommend small modifications to the survey design in the Tablelands regions based on the coverage assessment obtained using the recently available distance sampling survey design package in R (dssd; Marshall 2021). These will achieve more uniform coverage thereby ensuring that the surveyed area is a more representative sample from the study region.
- Increase the number of samplers (small areas, or blocks, containing zigzag transects) in the Western Plains and consider a randomised, systematic placement of blocks.
- Either include the areas of national parks in the survey design for the Western Plains or exclude them entirely (i.e., exclude the area when calculating the abundance estimates). Providing there are no constraints to surveying over national parks, transects falling within national parks could be surveyed and then the analyses later completed both excluding and including the data from the national parks, Alternatively, the national park data could be analysed separately.
- Continue to explore observer-specific detection functions to investigate potential differences in search behaviour between observers, such as guarding the line.
- Probabilistic approaches for matching duplicate detections could be explored for MRDS analyses of Western Plains detection data.
- For analysis of the Western Plains data, consider a wider pool of component models to choose from in model selection.
- Include uncertainty associated with any adjustment factors into the uncertainty of the density/abundance estimates so this better reflects the uncertainty associated with the total

abundance or density estimate. This will require estimates of variances (or coefficients of variation) for the adjustment factors.

- Consider updating the adjustment factors because these were based on data from surveys that took place over 20 years ago and the factors may have changed in this time.
- Continue to trial new technologies to evaluate their potential to deliver accurate and cost-effective abundance estimates. For example, the use of digital, visible light and infrared cameras could be trialled at the same time as conducting a helicopter survey to obtain a comparison. Long-range drones also have considerable advantages in terms of safety and cost compared to aircraft.

Contents

Introduction

The New South Wales (NSW) Department of Planning, and Environment has researched, monitored, and managed the NSW kangaroo population under the Kangaroo Management Programme (KMP) since 1974 (NSW KMP, 2021). Aerial surveys are undertaken annually, implementing distance sampling (Buckland *et al*., 2001) and mark-recapture distance sampling (Laake and Borchers, 2004) methods to estimate population density and abundance. The Centre of Ecology and Environmental Modelling (CREEM), University of St Andrews, Scotland, was approached to conduct a review of the survey design, protocols and analysis used by the KMP. This report contains the results of the review.

Background

There are four large species of macropod in NSW which are of particular interest because they are harvested for commercial purposes: red kangaroo (*Osphranter rufus*), eastern grey kangaroo (*Macropus giganteus*), western grey kangaroo (*Macropus fuliginosus*) and common wallaroo (*Macropus robustus*). Harvesting quotas for each species are based on population estimates and so it is important that these estimates are reliable and estimated with high precision. The red-necked wallaby (*Macropus rufogriseus*) and swamp wallaby (*Wallabia bicolor*) also occur in NSW and population estimates of these are also obtained.

The regions of interest for surveying in NSW are divided into the Western Plains and Tablelands. These are large regions containing a variety of habitats and terrain, and different survey methods are used in each of these regions.

The Tablelands region is subdivided into six kangaroo management zones (KMZs) and observers on helicopters implement conventional distance sampling (CDS) methods (Buckland *et al*., 2001) in these zones.

There are nine KMZs in the Western Plains; fixed-wing aircraft are used, and observers implement mark-recapture distance sampling methods (MRDS; Laake and Borchers, 2004) in these zones.

To provide some context to these surveys and assessment methods, CDS and MRDS are described in this introductory section, as well as the methods currently employed by the KMP.

Survey methods are improving all the time as new technologies develop and these offer the potential to obtain population estimates that are more accurate or with greater precision (Buckland *et al*., 2023). At the end of the report, new technologies appropriate for this application are discussed. Firstly, the scope of this review and the assumptions made are listed below.

Scope of the review

Terms of reference

Undertake a review of the methods and processes used to survey kangaroos across the commercial kangaroo harvest zones and provide the Department with a report for review and feedback. The report should include:

- Assessment of the appropriateness and adequacy of the design and execution of the NSW kangaroo population surveys (inland western NSW and tablelands commercial harvesting zones):
- Assessment and commentary on the sampling methods
- Assessment and commentary on the data collection methods
- Assessment and commentary on the robustness and accuracy of the data analysis methods

- Recommendations for improvements, including:
  - i. advice on quality assurance procedures in data collection, transfer, and analysis
  - ii. advice on survey design and conduct
  - iii. advice on the potential use of drones, thermal imaging and digital surveys in the future census of kangaroos in NSW.

Assumptions

The review was only concerned with the distance sampling methods that are currently being used, or may be considered in future, and not with methods used previously.

Consideration of alternative methods (e.g., census counts or the use of camera traps) was not part of this review.

It was assumed that the aim of the surveys is to estimate absolute density and abundance as accurately as possible for each year, rather than to estimate a trend over several years.

The analysis method to obtain harvest quotas was not part of this review although it was recognised that the abundance estimates obtained with the methods under review are used to obtain these quotas. Since these estimates are used to set harvest quotas, throughout this report we define a precautionary approach to be one which tends towards lower abundance estimates, rather than higher estimates.

Statistical methods

Two implementations of distance sampling methods are used in the surveys by the KMP: CDS (Buckland *et al.*, 2001) methods are used in the Tablelands and a related method, MRDS (Laake and Borchers 2004) is used in the Western Plains. These methods are briefly described below.

Conventional Distance sampling

In line transect distance sampling methods, multiple, randomly placed (or systematically placed with a random start point) lines, or transects, are superimposed over the region of interest. An observer (or team performing the same role) travels along the transect lines recording the perpendicular distances from the transect line to detected groups (here a group could be a single kangaroo or a group of kangaroos) either side of the transect. A proportion of groups may be missed (because, for example, groups are less visible far from the line). CDS uses the perpendicular distances to estimate the proportion detected, which in turn allows the density of individual animals (D) to be estimated as follows:

$$\widehat{D} = \frac{n\,\hat{E}[s]}{2wL\hat{P}_a}$$

where:

- n is the number of detected groups.
- $\hat{E}[s]$ is an estimate of the expected group size.
- w is the known strip half-width (and, with observers searching both sides of the transect, the total strip width is $2w$).
- L is the total length of the transects (and the total area covered by the survey is $2wL$) and
- $\hat{P}_a$ is the estimated proportion of groups detected within the covered area.

Estimated density is converted to estimated abundance ($\widehat{N}$) simply by multiplying by the area of the survey region (A):

$$\widehat{N} = A.\widehat{D}$$

Key information for each detected object is the perpendicular distance of the object from the transect and, if objects occur in groups, the number of objects in the group. The distances allow the proportion of detected objects to be estimated (see below) and the group sizes allow the expected group size to be estimated.

The uncertainty in the density estimate is a combination of the uncertainty in the encounter rate (n/L), expected group size and $\hat{P}_a$. It is usually expressed as a coefficient of variation (CV) and these components of variance are combined using the delta method (Seber, 1982). The variance of the encounter rate is estimated using a method developed by Innes *et al*. (2002) using the R2 form of the estimator as in Fewster *et al*. (2009).

Probability of detection

The proportion of detected objects is estimated from a detection function model, $g(x)$, fitted to the observed distribution of perpendicular distances, x, assuming all objects on the transect are detected (i.e., certain detection at perpendicular distance zero, denoted by $g(0)$=1).

Different forms of the detection function model might be appropriate depending on the shape of the distribution of perpendicular distances to detections. Two of the frequently used forms are the hazard rate or half-normal forms. Akaike's Information Criterion (AIC; Akaike, 1973) provides an objective quantitative measure for deciding between detection function forms. For reliable estimation of the detection function, Buckland *et al*. (2001) recommend that 60-80 objects are detected.

Multiple covariate distance sampling

In CDS (Buckland *et al*., 2001), the detection function is determined only by the perpendicular distances. In practice, there may be several detections functions, for example, for different observers or different terrain. The effects of covariates, other than distance, can be incorporated into the detection function model by setting the scale parameter in the selected detection function model to be an exponential function of the covariates (Marques and Buckland, 2004). Thus, the covariates can affect the rate at which detection probability decreases as a function of distance, but not the shape of the detection function. This is referred to as multiple covariate distance sampling (MCDS).

Assumptions of distance sampling

Distance sampling has some key assumptions that need to be met to obtain reliable density estimates and these are described below. There are other assumptions, but the ones listed below are the most important for this review.

A representative area of the study region is surveyed

The covered area of a survey is the area along the transects out to a truncation distance, w, either side of the transect (so the total area is $2wL$). The covered area should be a representative sample of the study region, and any point in the study region should have the same chance as any other point of being included in the survey; this property is referred to as uniform coverage.

Objects on the transect are detected with certainty

Objects on the transect are detected with certainty (frequently denoted by $g(0)$=1). If this not the case, the density estimates will on average be lower than the true density. The probability of detection on the transect can be less than one because objects are not available for detection (referred to as availability) or because observers miss objects on the transect (referred to as perception). To estimate perception, the probability of detection on the transect, MRDS methods can be implemented (see later). Availability is also discussed below.

Objects are detected at their initial location

Conceptually, DS methods require objects to be stationary as if a 'snapshot' was taken at the time of the survey and the distances of detected objects are measured perpendicular to the transect. This assumption can be violated by the animal moving before detection, and this may be in response to the observer or independent of the observer. As a guideline, Buckland *et al*. (2001) recommended that objects are moving no faster than half the speed of the observer. Nevertheless, distance measurement must be taken at the original location of the object; movement in response to the observer before being detected will result in P_a being over- or underestimated (depending on the nature of the movement).

Measurements are exact

Ideally, recorded perpendicular distances are exact (Buckland *et al*., 2001)). Sometimes this is not feasible, and distances are recorded in distance intervals, as is the case with the KMP surveys. In this case, careful measurement is required at the cut points between intervals. In addition, if groups of objects can span a large area, the distance to the centre of the group should determine the distance interval and all objects within the group should be included in the group size (Buckland *et al*., 2001, pp 272).

Detections are independent

Detections by observers are independent, hence, the detection of one object does not increase the likelihood of another detection.

Mark-recapture distance sampling

An MRDS approach can be implemented to estimate the probability of detection on the transect (Laake and Borchers, 2004). In these surveys, there are two observers (or observing teams) and there are several observer configurations which can be implemented depending on the situation (e.g., whether responsive movement is an issue). In aerial surveys (when the plane is moving quickly), responsive movement between detection by one observer and detection by the other, is unlikely to be a problem and in the KMP surveys in the Western Plains, two observers search independently (known as an independent observer (IO) configuration). As with conventional DS, the distances of detected objects are recorded but it is also important to be able to identify when the observers have detected the same object (called a duplicate).

The mark-recapture detection function defines the probability that an object at given perpendicular distance, x, with covariates z, was detected by observer q (q=1 or 2) given that it was seen by the other observer and is denoted by $p_{q|3-q}(x,z)$. It is modelled using the logistic form:

$$p_{q|3-q}(x,z) = \frac{exp\left(\beta_0 + \beta_1 x + \sum_{k=1}^{K} \beta_{k+1} z_k\right)}{1 + exp(\beta_0 + \beta_1 x + \sum_{k=1}^{K} \beta_{k+1} z_k)}$$

where β_0, β_1, ..., β_{k+1} represent the parameters to be estimated and K is the number of covariates other than distance. It is important to include all variables affecting detection probability in the model; unobserved, or unrecorded, variables that affect detection probability can mean that any correlation between the detection probabilities has not been accounted for (known as 'unmodelled heterogeneity). AIC can again be used for model selection. The intercept of the selected MR model provides an estimate of the probability of detection on the transect. If 'observer' is an important covariate, then the probability associated with one observer will not equal the probability associated the other observer.

The analyst has two choices when fitting models, a point independence (PI) model and a full independence (FI) model. Other options, e.g., limiting independence (Buckland, Laake and Borchers, 2010) exist but are not available in freely distributed software and so are not considered here.

PI models assume that it is only at distance x=0 (i.e., on the transect) that there is no unmodelled heterogeneity. To fit an IO point independence model, two subsidiary models are required: a distance sampling (DS) model fitted to all unique sightings, assuming that the intercept at perpendicular distance zero is one; and the MR detection function to estimate the probability of detection, by at least one observer, at distance zero. This probability of detection at perpendicular distance zero is then used to adjust the DS detection function to obtain an overall probability of detection.

The FI model assumes that there is no unmodelled heterogeneity at any distance and the MR model (as implemented in the R package mrds; Laake *et al*., 2022) is used to model the probability of detection on the transect and as distance increases. This model is simpler than the PI model (because only one model is fitted) but it can be difficult to be sure that the assumption holds.

Availability

Not all objects may be available to be detected (because, for example, they are hidden under trees) and this can result in abundance being underestimated if unavailable objects have not been accounted for. Availability of objects can result in underestimation in surveys with both single and double (teams of) observers. It is likely that in a fast-moving aircraft, if an animal is available to one observer it is likely to be available to the other and likewise if it is unavailable, it will be unavailable to both observers, hence, the estimate will be an estimate of available objects. If the proportion of available objects is known, or can be estimated, density estimates can be corrected to obtain estimates of all animals. Estimates of availability cannot, in general, be determined from the survey data but need to be estimated from other types of data, for example, by fitting telemetry tags to animals.

Implementation of methods

For the methods described above to result in reliable estimates, key assumptions need to be fulfilled. Key assumptions for DS were summarised by Glennie *et al*. (2015) as 1. animals on the line are detected with certainty; 2. measurements are exact; 3. animals are detected at their initial location. Careful design and implementation of the survey methods, training of observers and consistent application of survey protocols can help fulfil these assumptions.

The methods currently being employed by the KMP are briefly listed below.

Current surveys and analysis

Due to differing terrain and relief, different approaches to survey design and analysis are used in Western Plains and the Tablelands. Different species are found in these two regions.

Tablelands

The target species in this region are eastern grey kangaroos, common wallaroos, red-necked wallabies, and swamp wallabies.

Survey design

- The Tablelands is divided into three regions: Northern, Central and South-East Tablelands.
- One region is surveyed each year (and so each region is surveyed once every three years).
- The regions are divided into six KMZ: three KMZ in Northern Tablelands, two in Central Tablelands and one in South-East Tablelands.
- Each KMZ is stratified into low-, medium- and high-density strata, based on land capability to support kangaroos (for example, because of rainfall). Low strata are not surveyed, and other areas may also be excluded from surveying (e.g., open cut mine sites).
- Effort was calculated for each stratum individually to try and obtain a desired CV, of usually around 20%.
- A new set of transects are generated for each survey, with the effort required to achieve the target CVs based on the previous survey results from three years before.
- A stratified segmented grid or transect design was used, usually with transect segments being 5 km, 7.5 km, or 10 km in length.
- Samplers that are not wholly within a stratum are excluded.

Survey protocol

- Surveys are undertaken in the austral spring when temperatures are low, and kangaroos are likely to be active during daylight hours.
- A helicopter is used and flies at a constant speed of 93 km/h at a height of 61 m above ground.
- Flights last 2-3 hours after sunrise, or before sunset.
- One team of observers search for animals.

Data collection

- Observers voice-record their observations which are transcribed at the end of each session.
- Distance intervals are 0-20, 20-40, 40-70, 70-100 and 100-150 metres. These are marked via metal booms extending from the helicopter frame and these bins are offset to compensate for observers not being able to see directly below the helicopter.
- Groups of kangaroos spanning more than one distance interval were recorded as separate "observational" groups, one in each distance bin (with number of animals in each bin allocated accordingly).

Analysis

- Performed in Distance for Windows version 7.3 (Thomas *et al*., 2010)
- CDS and MCDS models are fitted, and AIC, along with other diagnostics, are used to select the final model.
- Separate models are obtained for each species and in each KMZ for eastern grey kangaroo (EGK), where sample size is adequate.
- 95% confidence intervals (CI) and CV are obtained by bootstrapping (999 replicates)
- The abundance estimate of common wallaroo is multiplied by 1.85 as it has been found that helicopter surveys of wallaroo underestimate numbers by approximately 1.85 (Clancy *et al*. (1997).

Western Plains

The target species for these surveys are red and grey kangaroos (and feral goats are also recorded).

Survey design

- The region is divided into nine KMZ of differing sizes.
- All KMZs are surveyed each year.
- A grid was randomly laid down over the whole region and blocks chosen at random from the grid.
- Blocks that were more than 50% within a national park, or where permission to survey could not be obtained, were excluded. This resulted in 56 blocks retained for sampling.
- Zigzag transects were laid down within each block using the same start point in relation to the grid block boundary. Transects were adapted to avoid crossing boundaries or national parks. The total length of transects in a block ranged from approximately 118 – 172 km.
- The same survey transects are used each year.

Survey protocol

- Surveys are undertaken in June-July when weather is cooler, and the target species were most likely to be active during daylight hours.
- Surveys start at first light because kangaroos are nocturnal, and temperatures are cooler.
- Fixed-winged aircraft were used and flown at a constant speed of 185 km/h and height of 76 m above ground.
- Observers were mostly volunteers, within the organisation, and many return each year and so there is some continuity of observers over time.
- Observers are trained in MRDS methods.
- There are a team of six, (including pilot and aerial observer) in the plane and an independent observer configuration is implemented.
- Observers rotate around seats and less experienced observers are not intentionally paired with experience observers.

Data collection

- Data are collected electronically, rather than using voice recorders, because it is more practical and avoids signalling a detection to the other observers.
- Each observer records their observations into a data logger via an Xbox controller.
- Perpendicular distances are collected in distance intervals (indicated by booms attached to wing struts) are 0-50, 50-100, 100-200 and 200-300 metres.
- The plane does not have a bubble window and so the distances intervals (above) are adjusted to account for observers not being able to view directly under the plane.

Analysis

- Performed in R (R Core Team, 2020) using mrds package (Laake *et al*., 2020).
- Different models are obtained for red and grey kangaroos.
- Assumed PI (so DS and MR component models are fitted) using AIC for model selection of covariates but also taking into account CV and domain knowledge of the detection function shape.
- The abundance of grey kangaroos is divided into western grey and eastern grey kangaroos using relative proportions determined by ground surveys.

Data and material provided for the review

Reports and data of recent surveys were made available to CREEM via a shared OneDrive folder. There were also several meetings between the two groups.

Overview of assessment

Distance sampling methods are regularly used to estimate animal abundance across large areas and are employed across a range of terrestrial mammal species, e.g., great apes (Ancrenaz *et al*., 2005), polar bears (Aars *et al*., 2017) and red deer (Trenkel *et al.*, 1997). Buckland *et al*. (2015) discussed the application of DS methods to surveys of ungulates stating that the methods may also be relevant to other large herbivores. Fewster and Pople (2008) demonstrated the suitability of the point independent (PI), mark recapture distance sampling models for aerial kangaroo surveys. These models are used to estimate animal abundance in the Western Plains.

Given the regions to be surveyed are particularly large, it would be impractical to cover them using ground-based surveys. Additionally, using aerial surveys more easily facilitates the random placement of transects and ensures that the observers in the planes or helicopters are moving at sufficient speed relative to the animals. Aerial surveys can also provide observers with a good forward view so they can record the position of detected animals before there is a chance of responsive movement.

Distance sampling methods rely on random transect placement and this has been a key component of the designs used for both the Tablelands and the Western Plains. The Tableland surveys used a randomly placed set of equally spaced parallel line segments and the Western Plains randomly selected blocks in which zigzag transects were generated. While the basis of both surveys was the random placement of transects, decisions, such as to only survey full line segments or to avoid national parks, will have meant that some parts of the survey region will have been less likely to be included in a survey than others. We make suggestions for refining the designs so that all areas of the study area have a more equal chance of being included in the survey.

While distance sampling does not require all animals within a strip to be counted, it does require that certain assumptions are met by the observers to accurately estimate the probability of detection within the strip. The observers in these surveys are provided with substantial training and aided in the task of accurate and efficient data recording using various technologies and, in the case of the Western Plains, dedicated software. However, some practical constraints and limitations have been noted on the selection of observers, with some being more experienced than others, and this has led to data analysis decisions to try to account for this.

The widely used software packages, Distance for Windows (Thomas *et al*., 2010) and the more recent R package mrds (Laake *et al*., 2023), are used to analyse the survey data to estimate density and abundance. Objective model selection methods are used to select detection function models but it is also recognised that certain models may be favoured to be consistent with what is known from experience of surveying kangaroos about the form that detection functions should take.

More specific details of all stages in the surveys are provided in the next section.

Details of assessment

The assessment is divided into sections for survey design, survey protocol, analysis, and new technology. Several topics could have been listed in more than one section but to avoid repetition such issues are dealt with in one section only.

Survey design

As a design-based method, distance sampling relies on valid survey design to generate reliable estimates of density and abundance and its associated variability. Firstly, the survey design should ensure that a representative sample of the study area is obtained. This must involve a component of randomness and every point in the survey region should have the same chance of being included in the covered area, namely uniform coverage. The consequence of unequal coverage will mean that some regions are less likely to be covered and, if these regions are different in some way to regions more likely to be covered in the survey, the resulting estimate will not be reliable estimates of density or abundance in the whole region. Analysis with designs with non-uniform coverage is possible but is more complex and requires coverage probabilities to be computed (Buckland *et al.*, 2001).

Secondly, sufficient independent samplers (e.g., transects) are required to reliably estimate the variability in encounter rate; usually this is a minimum of 10-20 samplers per stratum (Buckland *et al.*, 2001). Having sufficient samplers also helps ensure that individual surveys generated from a design provide a representative sample covering the different features (e.g., habitat types) present in the survey region. Systematically spaced designs also tend provide a more representative sample from the survey region than fully randomised designs.

The main considerations regarding survey design in both regions are summarised here. A more detailed discussion with coverage simulation results for the Tablelands is provided in Appendix A.

Tablelands

- The randomised segmented designs have achieved uniform coverage across the central parts of each stratum. In addition, the systematic nature of the segmented design ensures that any set of randomly generated transects will provide good coverage throughout the survey region for any individual set of transects.
- These segmented designs also provide over 20 independent samplers per stratum; sufficient to provide reliable estimates of the variability in encounter rate between samplers.
- There is some concern regarding areas of low coverage around the stratum boundaries and in small and disjointed areas. This is mostly due to only retaining samplers which fall completely within the stratum boundaries. If density is higher in these areas of lower coverage, then the abundance estimate will underestimate the true value and if it is lower, the abundance estimate will overestimate the true values. It is expected that density changes gradually from areas of higher to lower density and it is therefore possible that these regions of lower coverage on the stratum boundaries do differ in density than in the central parts of the strata, but this is not known to be the case. Coverage simulations can be found in Appendix A.

Western Plains

- This design aimed to achieve uniform coverage by laying down a 50x20 km grid at random across the entire study. A number of these 50x20 km blocks were then selected at random within each KMZ and equal spaced zigzag samplers were placed in the selected blocks. Within rectangular regions of the same size, an equal spaced zigzag design is known to give uniform coverage, as desired.
- Uniform coverage was somewhat compromised due to the exclusion of blocks which were comprised of more than 50% of national park. Selected blocks containing lower proportions of national park were sampled but transects were modified so that national parks were never surveyed. This will give no coverage in national parks and an area of lower coverage around the national park boundaries. The surface area of the survey region that was used for calculating the abundance estimates included the areas of national park, thereby effectively assuming that the density of animals within national parks is the same as that in

the rest of the study area. As these areas have not been surveyed for several years it would be difficult to establish whether this is a reasonable assumption.

- There are further areas that could not be surveyed due to landholder objection to low level flights. In some instances, modified transects to avoid such properties will have resulted in areas of higher coverage within the 50x20 km sampled blocks. Although these areas are small in relation to the stratum areas, they can comprise a significant part of the covered areas in each stratum and clipping the transects rather than modifying would be preferable (see Appendix A). The areas that could not be surveyed were later included in the calculations for total abundance. Given that all such areas cannot be established in advance (landowner permission is only sought after the surveys are designed) and assuming that there are no strong commonalities between the features of the properties, the current approach of assuming these are missing at random seems reasonable.
- Between four and seven blocks are sampled per management zone, leading to some concern over the reliability of the encounter rate variance estimates for each stratum (see Appendix D). Having low numbers of sampled blocks per stratum also means that there is a greater risk of over-representing some habitat types and under-representing others.
- It was noted that density surface modelling did not yield the same estimates of abundance as the mark recapture distance sampling. There are various reasons as to why this is the case (see Appendix A for further details).

Survey protocols

The methods implemented by the observers should ensure that the key assumptions of distance sampling are valid. These key assumptions are detection of objects on the transect is certain (or with MRDS methods this can be estimated), objects are detected prior to any responsive movement and accurate measurement of distances (or allocation to distance intervals).

Surveying from aircraft is often the only realistic option if the region of interest is large, as with the surveys here, and the animals can move quickly. Travelling by air also allows the random placement of transects which may be difficult if the terrain is such that following randomly placed transects on the ground is difficult, or impossible (Buckland *et al*., 2015).

Observer training

Observers are provided with training so that they know what to do to ensure the key assumptions of DS are valid. For the Western Plains surveys, observers can practice using the equipment and implement the data collection methods which can only improve their ability to record data. Specific comments are provided below.

- Observers are shown how to enter data into the aerial survey data logger. Fields 'Time', 'Observer', 'Position' are mandatory for data entry; fields 'Cloud', 'Temperature', 'Wind' are optional. It is unclear why any fields are optional. If this is useful information, e.g., for estimating the detection function, it should be mandatory to enter these fields.
- For each detection, the information recorded is limited to the most essential (species, group size, distance bin). This is useful given the short amount of time available for each entry in high density areas.
- Observers are trained to follow a specific search pattern (which was not obvious from the training document). However, it includes a key aspect of distance sampling surveys which is not to 'guard' the line.
- Observers are trained via specially designed video games to obtain counts of animals, distinguish between species, and assign groups to distance bins. For Western Plain surveys,

observers use these games to familiarise themselves with the Xbox controller used for data entry which requires memorising which button corresponds to which piece of information (species, numbers, distance bins). These games also expose the observer to situations during which they need to enter data very quickly, like surveying in high density areas.

- Observers are trained to obtain counts via subitising (i.e., identifying and recording small groups of 1-4 and, at least for Western Plain surveys, recording each of these via a press of a button on the Xbox controller. For groups larger than four, multiple buttons can be pressed within a single record. This is in accordance with Flemming and Tracey (2008) who recommended this method to avoid inaccuracies in data entry due to failure to remember correctly and reduce observer fatigue.

Distance measurement

- View of transect – the view directly below both the plane and helicopter is obscured and so the observers cannot see directly below the plane. To allow for this, the first distance bin has been offset and markers for cut points between distance bins (on booms attached to the frame) have been adjusted accordingly (according to a diagram in the training document). Hence, the assumption of certain detection on the line changes to assuming perfect detection at the offset distance from the line.
- Hilly terrain – hilly terrain can affect distance measurements because the plane/helicopter may not be at a consistent height above ground for all parts of the covered region. This is more of an issue in the Tablelands than the Western Plains and in the Tablelands transects are generally aligned perpendicular to contours to avoid land being at different heights on either side of the plane.

Data collection

- For Western Plains surveys, the detection data are entered into the data logger via an Xbox controller. This enables the observers to enter information without averting their eyes from the survey area. Specific keys are used for each of the species, groups of 1, 2, 3 or 4, and each of the distance bins. Each record starts with a species, followed by one or more subitised numbers and ends with a distance bin. Mixed groups (mobs) consist of multiple records whereby any second or further species are entered as separate records, however, the bin key is entered three times. A false entry can be marked as such by leaving out the bin entry as these records will be automatically deleted. Additional options are available for deleting false entries during flight.
- For Tablelands surveys, observers use voice recorders. Observers transcribe their recordings after each survey session.
- The data recording methods of the Western Plains surveys are in accordance with Flemming and Tracey (2008) who recommended to "record each observation onto tape, digital voice recorder or manual counter set up so that each unit from 1 to 4 is assigned to a different key. This prevents observers from looking away from the surveyed area to log sightings and eliminates the need to memorise numbers". Furthermore, recording multiple species increases observer alertness and ensures intensive searching for different search images.
- A concern with regards to responsive movement is that the view immediately to the front is obstructed as the plane does not have bubble windows. Animals avoiding the plane might be logged as they cross through one of the four bins while in sight. Information on movement of animals (e.g., direction in relation to plane, speed) might be useful to address this issue.

Definition of groups and allocation of groups to distances bins

- Group sizes can vary from one to five but can be much larger (e.g., 150 animals). How a group is defined (e.g., animals within a certain distance of each other) and how to determine the centroid of the group need to be clearly specified and applied consistently.
- There is a discrepancy between Tablelands and Westers Plains surveys in how observers are trained to define groups. For the former, observers separate a group that spans more than one distance bin into one group per bin. Western Plains observers are trained to record one group and attribute it to the bin which contains the centroid of the group. Here, a group is defined as an independent detection of multiple animals. More specifically, if detection of a first animal leads to detection of further animals that are nearby, observers are meant to call it a group. If a detection is independent from previously detected animals, it constitutes a new group.
- To avoid observers being tempted to record groups that are just beyond the furthest distance bin, observers are trained not to include groups that are beyond the furthest bin.

Identifying adult/juveniles

- All animals on the ground and large enough to be detected are included in counts. Joeys in pouch cannot be determined and so the survey is of adults and juveniles at foot (or animals large enough to be detected). Thus, the estimates may underestimate the true population unless some correction is made to account for young that are hidden. The surveys here do not attempt to include in-pouch young.

Mark-recapture distance sampling

MRDS methods are used to estimate the probability of detection on the transect by the observers.

- Having two observers/observer teams on the same plane will address the issue of observers missing animals that are in view on the transect but not the issue of animals not being available to be seen, due to being under vegetation cover, for example. Availability of animals does not appear to have been addressed specifically by any study and is discussed below, however, adjustment factors (which will likely include a component of availability) have been obtained from other studies.
- During observer training, Western Plains observers are given a short introduction into mark-recapture distance sampling. However, it is unclear whether the observers are made aware that it is crucial that they do not cue each other. Also, to be able to match detections, the data need to be entered as close in time/space to the actual detection as possible and the observers need to have clear instructions on what constitutes a group/cluster.

Identifying duplicates (Western Plains)

- Identifying duplicates is critical in MRDS analysis but can be difficult in practice, and this is recognised by the KMP in the observer training and in data preparation for analysis. It can be a particular problem in high density regions and if key information (such as group size, or distance bin) differs between observers.
- This issue is linked to the definition of what constitutes a group. For example, should one observer record six animals as one group and the other observer records the same animals as two groups of three, only one of the groups of three will be matched with the group of six as a duplicate.

- Fewster and Pople (2007) specified how they defined duplicates in a MRDS survey of eastern grey kangaroos. Their rules were deliberately designed to err on the side of overestimating the probability of detection on the transect, and, hence, underestimating the density.
- Currently, an algorithm is used by the KMP that matches duplicates by the following criteria: they need to be of the same species, within the same distance bin and within 200 m of each other (which corresponds to 4.5 seconds of flight time).
- There is interest from the KMP to switch to the methods described by Hamilton *et al*. (2017) who developed a probabilistic approach to determining duplicates; they used vertical angle and time difference between detections. The application of this approach to the surveys of interest here would require some adaptation because distance bins are recorded rather than exact downward angles.
- Methods for MRDS surveys which do not require duplicates to be identified have been developed recently (Stevenson *et al*., 2018; Borchers *et al*., 2022). They currently require detection probability not to depend on perpendicular distance, making them more suitable for surveys with cameras than with human observers. If surveys are done with cameras, use of these methods may well be worth investigating.

Analysis

Sample sizes (number of detected groups)

- For Western Plains surveys, the number of groups detected are much larger than minimum guidelines recommended by Buckland *et al*., (2001) i.e., 60-80 detections. This facilitates inclusion of covariates into the detection function for MCDS and MRDS (see below).
- For Tableland surveys, the number of groups detected for smaller species (i.e., wallaroos and wallabies) may not reach 60-80 within each KMZ. In these cases, data from different zones are combined to determine the detection function – this is a reasonable approach but assumes that detection probability is similar in different zones. If detection probability is different between the KMZs, abundance estimates in each KMZ will on average differ from the true abundance in the KMZ. Including zone as a covariate in the detection function can be tested (depending on the number of groups) to determine whether a term is required to account for different detection probabilities.
- Similarly, if sample size is very small, combining detections from different years, or strata, may result in a more appropriate detection function, if survey protocols, detectability, etc., can be assumed to be similar between years/zone. MCDS methods can then be used to determine if year (or zone) should be incorporated as a covariate in the detection function to account for any differences between years (or zone) (also see below).
- Sample sizes are reported for the Tablelands surveys but not, in general, for the Western Plains surveys. Reporting sample sizes would be helpful for readers to evaluate the basis on which the resulting abundance estimates were obtained.

Truncation

- When using exact distances, some truncation is often desirable to exclude outliers (Buckland *et al*., 2001). With distances recorded in distances bins, this is more problematic if there are only a few bins. Truncation will reduce the degrees of freedom used, for example, to perform Chi-square goodness of fit tests. However, if there are no groups, or a small number of groups, in the furthest bin(s), then truncation can improve model fit and this is especially pertinent if no covariates are included in the detection function. If covariates are included in the detection function, then truncation is not desirable because detections in the tails can inform model fit (see examples on Distance sampling.org).

Model selection

- As seen previously, the sample size (number of detected groups) is large for some species and provides the potential to investigate a large number of models. This is particularly important when fitting MRDS models to account for dependence between detections (Laake and Borchers, 2004). Appendix B illustrates a more extensive model selection search.
- AIC, along with goodness of fit statistics, provide an objective measure to assess detection function form and inclusion of covariates and is more defensible than, for example, considering CVs of the estimate.
- KMZ (or stratum) could also be used as a candidate covariate in the model. This would provide an alternative to fitting separate models to detections within each stratum, for example in the Central Tablelands, if sample sizes within strata are small.

Variance estimation

- With systematic survey designs (i.e., equally spaced samplers with a random start point), the samplers are frequently treated as independent for the purposes of calculating the variance. Fewster *et al*. (2009) showed that using this approach where there is a spatial trend in a population, can result in overestimating the variance. They provided alternative estimators which used poststratification (i.e., grouping adjacent samplers) to reduce the overestimation. In the Tablelands, a segmented grid design is used (where samplers are short and laid on a systematic grid) and so the variance estimate may be reduced with an adaptation of this approach. Fewster *et al*. (2009) derived an estimator for point transects which could also be used where lines are short, of equal length and positioned on a regular grid. If this was of interest, it would require analysis time because currently, this estimator is not available in Distance for Windows (Thomas *et al*., 2010) or in the distance sampling R packages and would require some adaptation for unequal length lines (which might occur if samplers are split where they cross a boundary). Further details are in Appendix D.
- In the Western Plains, using blocks rather than zigzag transects as sampling units has been discussed. The effect of doing this and suggestions are provided in Appendix D.

Availability of animals

- The availability of animals to be detected may affect surveys in both the Western Plains and Tablelands. Availability can be difficult to quantify but, without knowing/estimating the proportion of animals likely to be unavailable, any resulting abundance estimate may underestimate the true value if the proportion of unavailable animals is substantial and not accounted for. Availability is likely to be affected by habitat; for example, in open plains animals may be always available but in wooded areas animals may be hidden under the canopy.
- To maximise the time when animals are visible/available, surveys are undertaken at a time of year when temperatures are coolest and time of day when animals are most likely to be active (and not sheltering under trees), hence, the proportion of animals not being available to be detected is likely to be at a minimum.
- A basic methodology for determining availability would be to randomly choose groups of different known sizes and in different terrain/habitat and direct the plane to fly over them (at the same height and speed as employed during the surveys) and compare the observers'

data to the known data. It is recognised that this may be difficult to implement in practice and so an assessment of the likely gains of doing this would be required. In open habitats, the proportion of animals that are unavailable to be detected may be small and so concentrating effort to estimate availability in more tree-covered and wooded zones may be more worthwhile.

- There has been substantial work to assess correction factors between walked surveys and helicopter and fixed-wing aerial surveys (McLeod, unpublished). It is likely that these correction factors include a component due to availability of animals because, for example, animals under trees are more available to observers undertaking ground surveys. The correction factor for common wallaroos is discussed below and in Appendix C.

Probability of detection on the transect

- MRDS methods are used to estimate the probability of detection on the transect. If this probability is less than one, the abundance estimate will underestimate the true value on average.
- These methods are implemented during the surveys of the Western Plains, and so the probability of detection on the transect can be incorporated into the overall detection probability (see Appendix B).
- For the analysis of surveys in the Tablelands, detection on the transect is assumed to be certain. If this is not the case, then the resulting estimates will be an underestimate of the true values. To account for this in estimates of common wallaroo, a correction factor is applied.
- A comparison of estimates from helicopter line transect surveys and walked line transect surveys by Clancy *et al*. (1997) provided an estimate of perception on the transect and availability. They found estimates for red and eastern grey kangaroos did not differ between these two methods, but for common wallaroos the difference was substantial; the density estimated from helicopter surveys was 54% of the density estimated from walked surveys, thus, a correction of 1.85 (=1/0.54) is applied to estimates of wallaroo (see later). Undertaking a double-observer survey on the helicopter surveys would allow the probability of detection on the transect to be estimated – this observer configuration could be employed for a portion of the survey such that enough detections are generated for fitting a MR model, However, this does depend on being able to accommodate a double-observer configuration in the helicopter. Correction factors associated with other species, accounting for habitat, are summarised in McLeod (unpublished) but are not applied.

Movement of kangaroos

- Distance sampling methods are essentially ‘snapshots’ of animals in the covered region but kangaroos are mobile animals. Red kangaroos can move at up to 70 km/h for short distances[1] with the maximum speed for smaller species being slower (e.g., 65 km/h for eastern grey kangaroos[2] and 48 km/h for wallabies[3]).
- Movement of animals during surveys can be either in response to the survey or independent of the survey.
    - Responsive movement

[1] https://www.speedofanimals.com/animals/kangaroo
[2] https://www.wildlife.vic.gov.au/__data/assets/pdf_file/0013/114034/Eastern-Grey-Kangaroo.pdf
[3] https://z-upload.facebook.com/wwfaustralia/posts/10162273281728712

        - Responsive movement of kangaroos during surveys is considered by observers to be small.
        - Aerial surveys can be an advantage when there is responsive movement because observers can look far forward and detect animals before they have responded. A concern with regards to responsive movement is that the view immediately to the front is obstructed as the plane does not have bubble windows and less experienced observers maybe looking more in the second bin (see Appendix B) or waiting before recording the distance bin. Animals avoiding the plane might be logged as they cross through one of the four bins while in sight.
        - Information on movement of animals (e.g., direction in relation to plane, speed) might be useful to address this issue.
        - Allowing observers to record the distance bin first, before other fields, may reduce the problem with responsive movement.

    - Movement independent of the survey
        - Glennie *et al*. (2015) quantified the relative difference between average estimated abundance and the true value due to independent movement in line transect surveys. The reported difference was around 10% for animals moving at 80% of the speed of the observer. Red kangaroos can potentially move at 40% of the speed of the plane and eastern grey kangaroos at 70% the speed of the helicopter. Glennie *et al*. (2021) showed that for these speeds, the relative difference was within ±5% of the true density.
        - Glennie *et al*. (2015) provided advice for reducing these differences e.g., searching further to the side and a shorter distance ahead. They realised that this contradicts advice for dealing with responsive movement because the observers are not searching far ahead. They concluded that for aerial surveys, issues due to independent movement were likely to be small. If this is of concern, or to quantify the effect for the specific conditions of the surveys under consideration, Glennie *et al*. (2021) provided a framework for incorporating animal movement into distance sampling methods. Additional information is required (e.g., from fitting telemetry tags to a sample of kangaroos) and this would need to be collected at a similar time of year as the surveys and in a similar terrain. An R package has been developed to assist with the analysis (i.e., moveds; Glennie *et al*., 2021).

Adjustment factors

- Factors are applied to divide the estimate of grey kangaroos into eastern and western grey kangaroos, and to adjust the estimate of common wallaroos.
- These factors are based on studies conducted over 20 years ago and so consideration should be given to updating these factors.
- The uncertainty associated with these factors is currently not being incorporated into the uncertainty and this should be done to better reflect the uncertainty associated with the density estimate. Incorporating the factor uncertainty can be done using the delta method (Seber, 1982) if it can be assumed that the components in the equation are independent. More details are provided in Appendix C.

New technologies

Wildlife monitoring is embracing new technologies and the use of drones to replace the helicopter and plane surveys and digital (e.g., video) cameras replacing observers, to record animals were discussed during meetings between CREEM and the KMP. Buckland *et al*. (2023) outlined the advantages of both these developments and other studies show promising results for using digital or thermal imagery (e.g., Lethbridge *et al*., 2019; Grierson and Gammon, 2002).

Detector types

- To investigate digital surveys further, we recommend conducting a trial survey comparing visual surveys simultaneously conducted with digital imagery and thermal imaging of the same survey strip. A similar approach has been conducted by Grierson and Gammon (2002) who pointed out the strengths of the different observation methods. However, that study did not consider different species. Such a trial survey might allow using the strength of each detection method: the acuity and ability to detect motion by human observers, the high-resolution imagery available for digital surveys nowadays (and the ability to use image analysis for detection) and the ability to detect camouflaged animals by contrasting body temperature in thermal imagery. A combination of these three approaches in a trial survey might allow the assessment of the percentage of animals missed by each method.

Drones

Drone-borne cameras have the potential to provide a safe and efficient way to sample and to estimate abundance, at lower cost. Issues to consider in using drones with cameras, rather than human observers in aircraft, include the following:

- Fixed-wing drones will be required to have adequate range for the surveys. Rotary-wing drones will not have adequate range although rotary-wing drones may have a useful role to play in species identification.

- Drone-borne cameras will have substantially narrower fields of view than human observers, resulting in narrower surveyed strips and lower sample size per distance flown. Strip widths can be increased by flying higher but increasing flight altitude reduces the resolution of camera images and reduces the ability to identify target species in images (whether identification is done by humans or automated identifiers). Although drone-borne cameras survey narrower strips, they could potentially deliver similar, or even higher, sample size per unit cost than human-carrying aircraft because they are substantially cheaper per distance flown.

- Human observers in the field can be better at detecting target species than automated identification methods (or human observers) are at detecting target species in video or still camera footage. While this may change in future, we suggest investigating the detection efficiency of automated detection before committing to large surveys using digital cameras, and to inform design decisions about such surveys.

- With relatively narrow angles of view, drone-borne (or other) cameras typically have little or no drop-off in image resolution from the centre to the edges of images. Consequently, there is no perpendicular distance detection function to model, simplifying analysis methods somewhat. In addition, there are analytic methods designed for estimating abundance from double-observer (or double-camera) survey data of this sort without the need to assume that detection is certain within images, and without the need to identify duplicates (see

Stevenson *et al.*, 2018 and Borchers *et al*., 2022). We suggest investigating the feasibility of these methods for drone-borne camera surveys.

Recommendations

Recommendations raised during this review are listed below. Some recommendations may be relatively straightforward to implement with a small cost implication (for example, additional time required for the analyst), while others have substantial associated costs (e.g., commissioning more surveys) and so these are listed as aspects for consideration by the KMP.

Some of the recommendations are provided for one set of surveys (or subsequent analysis) but not for another and so not everything listed will be applicable to both surveys (and analyses).

Survey design

Tablelands

- The segmented grid design should be considered generally preferable to the segmented trackline design; the latter can cause less uniform coverage in complex areas. The exception to this is when regularly spaced landscape features are present (e.g., fields of regular size) and the spacing of the segmented grid is like the regular pattern, but this is not thought to be the case in the Tablelands KMZs.
- Continue to use the same spacing both between the segment ends and between each row of segments because this ensures that each segment can be treated as an independent sampler.
- Use a design with equal coverage across strata within KMZs, i.e., use the same segment lengths and spacings. Keeping coverage the same across strata will ensure that when detections are pooled across strata for estimating the detection function, pooling robustness applies, and the detection function is representative of the KMZ as a whole (Buckland *et al.* 2001). See Appendix A for further discussion.
- To obtain a more uniform design coverage, it is recommended to generate the survey across the whole management zone and then later divide the samplers by stratum. This will remove the areas of lower coverage on the internal stratum boundaries. It is also recommended to consider a lower threshold for retaining segments, for example retaining all segments more than 50% inside the study area will greatly improve the coverage, see Appendix A.
- It is anticipated that the costs of implementing these changes will be minimal since only minor changes are recommended but it may add to the cost of post processing the data for analysis.

Western Plains

- If abundance estimates are to include areas of national parks, then these regions should also be included in the survey design. Otherwise, the abundance estimates should be calculated with the areas of national park excluded from the total area.
- Consideration should be given to increasing the number of blocks sampled in each KMZ to achieve a more reliable estimate of encounter rate variability. This change could involve sampling a greater number of smaller blocks This approach will also increase the chances of obtaining a representative sample. With more samplers per KMZ, the sampled habitat may stand a greater chance of being more representative of the habitat throughout the KMZ. It would, however, increase the time and cost of the surveys.

- If sampling blocks are to be redefined, a systematic design across the KMZ will generally increase the chances of obtaining a more representative sample.
- To better preserve uniform coverage, consider clipping the zigzag transects around areas that cannot be covered rather than adjusting the angles of the zigzag lines (see Appendix A).

Data collection

- To reduce the possibility of responsive movement, a suggestion is to reconfigure the digital data entry to allow observers to record distance bin as the first entry.
- The benefits of this would need to be weighed up with the ease of implementing the change and changing established data entry routines for returning observers.

Analysis

The recommendations in this section are likely to require additional time for the analyst to implement and investigate new approaches.

- In Western Plains surveys, use block, rather than each zigzag transect, as the sampler to estimate encounter rate variance. Increasing the number of blocks is recommended above.
- In Tablelands surveys investigate the use of the systematic variance estimators (e.g., Fewster *et al*. 2009).
- In the Western Plains surveys report, provide a summary of survey data e.g., number of groups, number of duplicates, by KMZ. This would allow readers to assess sample sizes.
- Consider expanding the list of component models for model selection. Model selection can be done separately for DS and MR models.
- Evaluate how a probabilistic approach to duplicate matching according to Hamilton *et al*. (2018) compares to the fixed tolerance approach currently used. Several challenges would need to be overcome. The component for calculating the score in Hamilton *et al*. (2018) pertaining to downward angle would need to be replaced with an equivalent term pertaining to distance bin. Furthermore, common dolphins investigated by Hamilton *et al*. (2018) have very different grouping behaviour compared to kangaroos or other species of interest here. Hence, it is uncertain if the method would work similarly well.

Adjustment factors

- Update, or review the applicability, of the adjustment factors as these are based on surveys which took place over 20 years ago and the proportions may have changed during this period. This would provide confidence that appropriate correction factors were being applied but would require substantial investment, hence, more detailed knowledge is required.
- Include uncertainty of adjustment factors into uncertainty of density/abundance estimates to better reflect the uncertainty of the density/abundance estimates.

New technologies

- To investigate digital surveys further, we recommend conducting a trial survey comparing visual surveys simultaneously conducted with digital imagery and thermal imaging of the same survey strip. A combination of these three approaches in a trial survey might allow assessment of what percentage of animals were missed by each method.
- We suggest investigating the detection efficiency of automated detection before committing to large surveys using digital cameras, and to inform design decisions about such surveys.

- We suggest investigating the feasibility of analytic methods for drone-borne camera surveys.
- These have important advantages in terms of pilot and observer safety but would require substantial investment to evaluate.

References

Aars J, Marques TA, Lone K, Andersen M, Wiig Ø, Bardalen Fløystad IM, Hagen SB and Buckland ST (2017) The number and distribution of polar bears in the western Barents Sea. *Polar Research* 36, 1374125. https://doi.org/10.1080/17518369.2017.1374125

Akaike H (1973) Information theory and as extension of the maximum likelihood principle. In International Symposium on Information Theory, 2nd edition (eds. Petran BN and Csàaki), Akadèemiai Kiadi, Budapest, Hungary, pp 267-81

Ancrenaz M, Gimenez O, Ambu L, Ancrenaz K, Andau P, Goossens B, et al. (2005) Aerial Surveys Give New Estimates for Orangutans in Sabah, Malaysia. *PLoS Biol* 3(1): e3. https://doi.org/10.1371/journal.pbio.0030003

Borchers DL, Nightingale P, Stevenson BC and Fewster RM (2022) A latent capture history model for digital aerial surveys. *Biometrics* 78: 274–285. https://doi-org.ezproxy.st-andrews.ac.uk/10.1111/biom.13403

Bouchet P, Miller DL, Jason J. Roberts JJ, Mannocci L, Harris CM and Thomas L (2020) dsmextra: Extrapolation assessment tools for density surfcae models. Methods in Ecology and Evolution. DOI: 10.1111/2041-210X.13469

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL and Thomas L (2001) *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press, Oxford, UK

Buckland ST, Borchers DL, Marques TA and Fewster RM (2023) Wildlife population assessment: Changing priorities driven by technological advances. *Journal of Statistical Theory and Practice* 17 (20) https://doi.org/10.1007/s42519-023-00319-6

Buckland ST, Rexstad EA, Marques TA and Oedekoven CS (2015) *Distance sampling: Methods and Applications*, Springer

Cairns S and Gilroy J (2001) Re-appraisal and enhancement of the current methodology used in the estimation of kangaroo populations in western New South Wales. Unpublished report to West 2000.

Cairns SC, Pople AR and Gilroy J (unpublished data)

Caughley G, Brown B, Dostine P and Grice D (1984). The grey kangaroo overlap zone. Australian Wildlife Research 11: 1-10

Clancy TF, Pople AR and Gibson (1997) Comparison of helicopter line transects with walked line transects for estimating densities of kangaroos. Wildlife Research 24: 397-409

Curtis E and McLeod S (2021) 2021 Western Plains Aerial Kangaroo Survey Results. Unpublished report DPE KMP

Curtis E and McLeod S (2022) 2022 Western Plains Aerial Kangaroo Survey Results. Unpublished report DPE KMP

Fewster RM, Buckland ST, Burnham KP, Borchers DL, Jupp PE, Laake JL and Thomas L (2009) Estimating the encounter rate variance in distance sampling. *Biometrics* 65: 225-236

Fewster RM and Pople AR (2007) A comparison of mark-recapture distance-sampling methods applied to aerial surveys of eastern grey kangaroos. *Wildlife Research* 35(4): 320-330 https://doi.org/10.1071/WR07078

Fleming PJS and Tracey JP (2008) Some human, aircraft and animal factors affecting aerial surveys: how to enumerate animals from the air. *Wildlife Research* 35(4): 258-267

Glennie R, Buckland ST and Thomas L (2015) The effect of animal movement on line transect estimates of abundance. *PLoS ONE* 10(3): e0121333. doi:10.1371/journal.pone.0121333

Glennie R, Buckland ST, Langrock R, Gerrodette T, Balance LT, Chivers SJ and Scott MD (2021) Incorporating animal movement into distance sampling. *Journal of the American Statistical Association* 116:533, 107-115, DOI: 10.1080/01621459.2020.1764362

Grierson IT and Gammon JA (2002) The use of aerial digital imagery for kangaroo monitoring. *Geocarto International* 17:2, 45-52, DOI: 10.1080/10106040208542234

Hamilton, ONP, Kincaid, SE, Constantine, R, Kozmian-Ledward, L, Walker, CG, Fewster, RM (2018) Accounting for uncertainty in duplicate identification and group size judgements in mark–recapture distance sampling. *Methods in Ecology and Evolution* 9: 354– 362 https://doi.org/10.1111/2041-210X.12895

Innes S, Heide-Jørgensen M-P, Laake JL, Laidre KL, Cleator HJ, Richard P and Stewart REA (2002) Surveys of belugas and narwhals in the Canadian High Arctic in 1996. *NAMMCO Scientific Publication* Vol. 4: 169-190 https://doi.org/10.7557/3.2843

Laake JL and Borchers DL (2004) Methods for incomplete detection at zero distance. In Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL and Thomas L (eds) *Advanced distance sampling*. Oxford University Press, Oxford, UK

Laake J, Borchers D, Thomas L, Miller D, Bishop J (2023) mrds: Mark-Recapture Distance Sampling. R package version 2.2.8, https://CRAN.R-project.org/package=mrds.

Lethbridge M, Stead M, Wells C (2019) Estimating kangaroo density by aerial survey: a comparison of thermal cameras with human observers. *Wildlife Research* 46: 639-648

Marques FFC and Buckland ST (2004) Covariate models for the detection function. In Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL and Thomas L (eds) *Advanced distance sampling*. Oxford University Press, Oxford, UK

Marshall L (2021) dssd: Distance sampling survey design. R package version 0.3.0 https://CRAN.R-project.org/packages=dssd

McLeod SR (unpublished) Correction factors used for estimating the abundance of kangaroos by fixed-wing aerial survey.

NSW Kangaroo Management Program (2021) Kangaroo population monitoring methods, Department of Planning and Environment.

Seber GAF (1982) *The Estimation of Animal Abundance and Related Parameters*. Macmillian, New York

Stevenson BC, Borchers DL and Fewster RM (2018) Cluster capture-recapture to account for identification uncertainty on aerial surveys of animal populations. *Biometrics* 75: 326-336. https://doi.org/10.1111/biom.12983

Thomas L, Buckland ST, Rexstad E, Laake JL, Strindberg S, Hedley SL, Bishop JRB, Marques TA and Burnham KP (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology* 47: 5-14

Trenkel VM, Buckland ST, McLean C and Elston DA (1997) Evaluation of aerial line transect methodology for estimating red deer (*Cervus elaphus*) abundance in Scotland. *Journal of Environmental Management* 50:39-50. https://doi.org/10.1006/jema.1996.0094. https://doi.org/10.1006/jema.1996.0094

Appendices

The appendices contain further details on:

Appendix A: Survey designs

This appendix expands the discussion on points raised in the main report in relation to survey design. It provides clarification on edge effects for line transect surveys and discusses the use of stratification. Coverage simulations are detailed for an example management zone from the central Tablelands region and further discussion of the likely coverage of the Western Plains design is included. R code (R Core Team, 2020) is provided to illustrate how coverage for designs can be assessed. This appendix concludes with a description of investigations into the spatial modelling of the Western Plains region in an attempt to ensure that a representative sample of the study area has been obtained in the surveys.

Edge Effects

The surveys in both the Western Plains and the Tablelands regions both use line transects designs. It is usual when performing line transect surveys to implement the more efficient minus sampling approach. Minus sampling means that nothing more is surveyed once the transect crosses the survey region boundary. Therefore, unless all transects meet completely straight stratum boundaries at an exact perpendicular angle then this approach will lead to a narrow area of lower coverage around the stratum boundary, see Figure A1 below. These standard, line transect, edge effects are very rarely severe enough to have significant impact on the estimates. Susceptible regions have complex shapes and long boundary lengths in comparison to the area they contain (Buckland *et al.*, 2001). In addition, having small truncation distances compared to the scale of the study region, as in these surveys, also helps minimise these effects. As such, these basic edge effects are not considered to be problematic for these surveys in neither the Western Plains nor the Tablelands surveys.

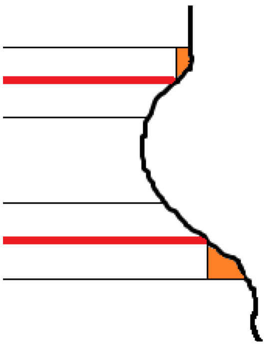


Figure A1. Illustration of areas not covered in line transect minus sampling design. Thick black line - survey region boundary, thick red line - transect, thin black lines - parallel to transect at the truncation distance, orange fill - areas not covered due to minus sampling.

While the basic edge effects are not considered problematic, the Tablelands surveys use segmented line transect designs which can cause more severe edge effects. The segmented line designs allow users to specify a threshold to avoid travelling large distances to survey very small lengths of transect. In the Tablelands surveys, the protocol was to only survey segments that were fully within the stratum boundaries. This leads to areas of lower coverage around the stratum boundaries which are several kilometres wide and represents a substantial proportion of the survey region. Coverage simulations presented in this section investigate the coverage of a design similar to current practice and also investigate how coverage can be improved to be more uniform across the study region.

Stratification

Stratification can be employed for several purposes. Two common reasons are 1) to generate estimates of abundance at the stratum level and 2) to improve precision in the overall estimate of abundance for the whole study region. The Western Plains region employs stratification to obtain estimates for each KMZ separately. The Tablelands surveys also stratify by KMZ but also further stratify by expected kangaroo density into areas of high, medium and low density.

When stratifying, there are various ways to allocate effort to each stratum. The safest approach, especially for multi-species surveys, is to use equal coverage in all strata. This would ensure that if data needed to be pooled across strata, e.g., for less abundant species, then accurate estimates of abundance would still be obtained due to pooling robustness (Buckland *et al*., 2001). If the strata have been defined accurately in relation to animal density, even if the same coverage is used in each stratum, the effect will be to improve precision in the overall estimate of abundance. This occurs because the variability in encounter rate between transects within each stratum should now be lower than if all transects were contained within the same stratum, and instead of contributing to the overall variability, the differences in encounter-rate between strata is now modelled explicitly (Buckland *et al*., 2001).

There are other more complex methods of allocating effort to strata depending on the survey objectives. If estimates are required at the stratum level, or to be able to compare density between strata, then effort may be allocated to obtain a target coefficient of variation (CV) for each stratum. In this case, the required effort can be calculated individually for each stratum based on the target CV. This approach will allocate higher coverage to where there are fewer animals and lower coverage to where there are higher densities of animals. In contrast, if the objective is to achieve the most precise estimates over the whole study area for all strata combined, then an alternative approach is implemented. Now we wish to allocate higher coverage to areas of higher density and lower coverage in areas of lower density, the opposite to the earlier approach. Details of both these options can be found in section 7.2.2.3 of Buckland *et al*., 2001. These, more complex, approaches need to be accounted for at the analysis stage and rely on accurate prior knowledge of animal distribution.

Tablelands

Existing Design

The Tablelands region is divided into three main regions: Northern, Central and South-East. The surveys are conducted in only one of these regions in each year, meaning that each is surveyed once every three years. Within these three regions there are six management zones, three in Northern region, two in Central and one in the South-East. Each of these management zones are then further divided into strata relating to areas high, medium and low density. The low density regions are not surveyed and do not contribute to the total estimated abundance for the management zones. Effort was assigned to the medium and high density strata to achieve a target CV of around 20% in each. This led to different coverage in the two strata with the medium density stratum having slightly higher coverage than the high density stratum.

The Central Tablelands south management zone is provided as an example to assess the coverage of the designs (Figure A2). The low density stratum is not included in the design, or estimation, so has been excluded in this example.

```
region.CTS <- make.region(region.name = "Central Tablelands - South",
                strata.name = c("High", "Medium"),
                units = 'm', shape = "shapefiles/CTS_Strata.shp" )

plot(region.CTS)
```

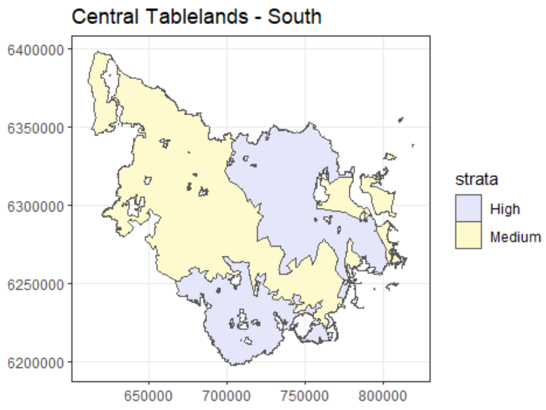


Figure A2. Study region for Central Tablelands south management zone. The area is stratified by expected animal density: high and medium.

Surveys are usually generated within Distance for Windows (Thomas *et al*., 2010) using either a segmented grid or a segmented trackline design in each stratum. Different length segments of 5km, 7.5km and 10km are investigated and usually segments of 7.5km are selected. Only segments that fall completely within each stratum are retained. Randomised line segments are generated for each survey based on the selected designs. Both segmented designs give uniform coverage in the central parts of the strata, however, excluding incomplete samplers causes lower coverage around the stratum edges. The coverage of each design was assessed as part of the survey design process in Distance for Windows (Thomas *et al*., 2010) but, due to the plot options and display colours, differences in coverage were less obvious than when investigated in R using the distance sampling survey design package, dssd (Marshall, 2021).

An example design for the Central Tablelands south management zone is defined below based on the designs currently used in the Tablelands KMZs. The spacings for each stratum were selected based on the target CVs for each of the strata and the encounter rates from the previous survey of the study area. A segment length of 7.5 km (7,500 m) was used and only segments which fell entirely within each of the strata were retained, i.e., a segment threshold of 100%. The design angle was chosen to be consistent with previous surveys. A set of example transects from this design can be seen in Figure A3.

```
# load a coverage grid from file as it takes time to construct
load("shapefiles/cover_CTS.robj")

# Define a design similar to that used in current surveys
design.CTS <- make.design(region = region.CTS,
                transect.type = "line",
                design = "segmentedgrid",
                design.angle = 90,
                spacing = c(10700, 9900),
                seg.length = 7500,
                seg.threshold = 100,
```

```
                 edge.protocol = "minus",
                 truncation = 150,
                 coverage.grid = cover.CTS)

# Generate a set of example transects
eg.transects <- generate.transects(design.CTS)
plot(region.CTS, eg.transects)
```

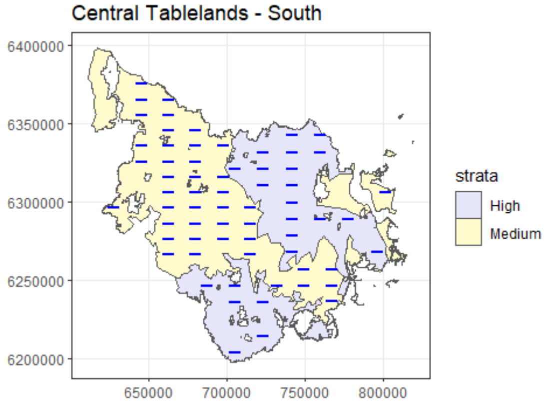


Figure A3. Example transects from a design similar to the design currently used in the surveys.

To assess the coverage of the currently used design, a coverage simulation was run. The coverage simulation involved generating 9,999 sets of transects from the design defined above. The design included a coverage grid with approximately 3,000 points spaced out throughout the study region in a grid layout. The coverage simulation records the proportion of times each of the coverage grid points are included within the covered areas of each of the 9,999 randomly generated surveys.

Figure A4 shows the coverage scores across the Central Tablelands south management zone. The combination of complex stratum boundaries and only retaining complete samplers leads to significant areas of lower coverage and some parts of the study region have no chance of being included in a survey. If density in these regions differs from the rest of the study area, then abundance would either be under or over estimated.

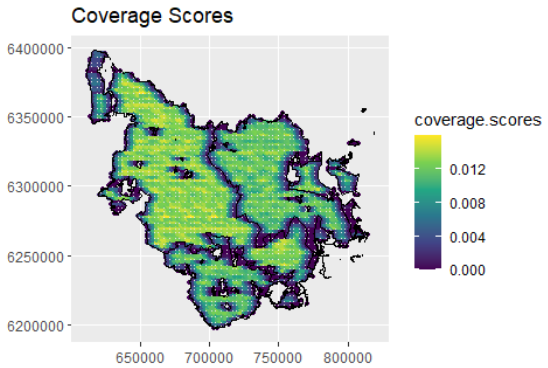


Figure A4. Coverage scores across the Central Tablelands south management zone for the existing unequal coverage stratified design.

The histogram of the coverage scores shows that a significant proportion of the survey region has no chance of being included in the surveys (Figure A5). Almost 9% of the management zone has a coverage score of 0, 30% has a score of less than 0.005 and 42% has a score less than 0.009. The values of 0.009 and 0.005 have been chosen arbitrarily to represent what could be classed as low and very low coverage values, respectively, in comparison to what would be expected under a typical uniform coverage design. As these are coverage simulations, there will be some stochasticity in coverage scores, but we would expect these to be roughly symmetrically distributed around a central value. For this example, that central value is expected to lie around 0.0125 based on the upper peak of the values in Figure A5.

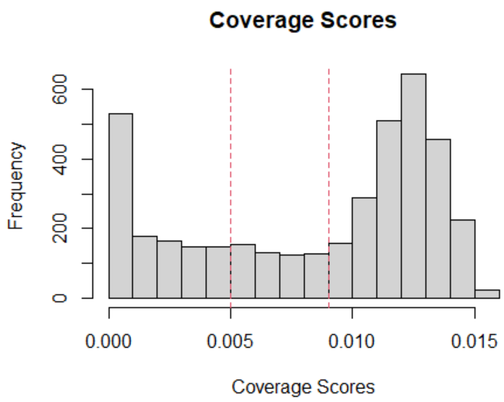


Figure A5. Histogram of the coverage scores across the Central Tablelands south management zone for the existing unequal coverage stratified design. The arbitrarily chosen values of 0.005 and 0.009 are indicated by red dashed lines.

Assessing alternative designs

Global design with complete samplers

The edge effects caused by only retaining complete samplers are amplified by the complexity of the internal stratum boundaries. Firstly, we propose generating an equal coverage design across the whole management zone. The samplers may still be allocated to individual stratum at a later stage. The aim of this would be to remove the edge effects along the internal stratum boundaries but retain the benefits of stratifying the design into areas of differing expected animal density. Stratifying the region by expected animal density, as has been done previously, can improve the overall precision of the estimates by reducing the variability in encounter rate between samplers within each stratum. In addition, the move to equal coverage across strata could also improve precision of the overall estimates as we should now be allocating higher coverage to where there are expected to be higher densities of animals. In practice, maintaining equal coverage across strata is usually the preferable solution, especially for mixed species surveys when different species may have different distributions. Having the same coverage in both strata also means that pooling robustness applies, and detections of less abundant species can then safely be pooled together across strata.

Below we define a modified design which generates a similar segmented grid design but over the whole management rather than by each stratum separately. An intermediate spacing value was selected, and the segment threshold was kept at 100, meaning that only segments which fall entirely within the study area (i.e., 100%) are retained. Examples transects from this design are shown in Figure A6.

```
# Read in the stratified area
region.shape <- sf::st_read("shapefiles/CTS_Strata.shp", quiet = TRUE)
# Create a global area by creating a union
global.shape <- sf::st_union(region.shape)


# Make a global region
global.CTS <- make.region(region.name = "Central Tablelands - South",
                          units = 'm',
                          shape = global.shape)

# Define the new design
design.CTS.100T <- make.design(region = global.CTS,
                          transect.type = "line",
                          design = "segmentedgrid",
                          design.angle = 90,
                          spacing = 10500,
                          seg.length = 7500,
                          seg.threshold = 100,
                          edge.protocol = "minus",
                          truncation = 150,
                          coverage.grid = cover.CTS)

# Generate some example transects
eg.transects <- generate.transects(design.CTS.100T)
# We can plot these transects against the original stratified
plot(region.CTS, eg.transects)
```

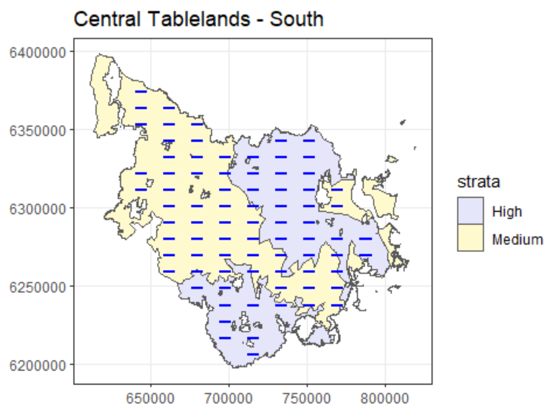


Figure A6. Example transects from a design which generates transects using a global region using a segment threshold of 100.

The coverage scores from the coverage simulation for this design are plotted in Figure A7. The coverage is now more uniform across the majority of the region, however, because only samplers that fall entirely inside the study region are retained, there are still significant edge effects around the exterior of the management zone and some of the smaller disjointed areas have very low or no coverage.

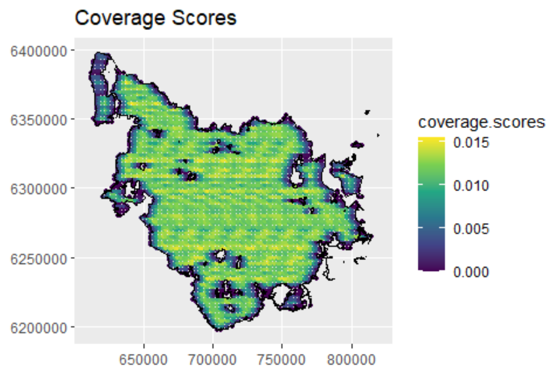


Figure A7. Coverage scores across the Central Tablelands south management zone for the global design using a segment threshold of 100.

A histogram of the coverage scores for this design still indicates several coverage grid points have no chance of being included in the survey (Figure A8). For this design, 4.2% of the coverage grid points have coverage scores of 0, 19.0% have coverage scores of less than 0.005 and 29.3% have coverage scores less than 0.009.

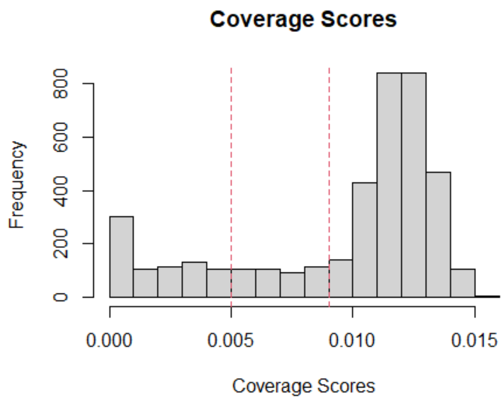


Figure A8. Histogram of the coverage scores across the Central Tablelands south management zone for the global design using a segment threshold of 100. The arbitrarily chosen values of 0.005 and 0.009 are indicated by red dashed lines.

Global design allowing split samplers

One way to further improve the coverage of this design would be to allow split samplers. In this example, we will retain any sampler that falls more than 50% inside the study area by setting the segment threshold to 50. All other design parameters remain the same as the previous example. Example transects from this design are shown in Figure A9.

```
# Define the design with a 50% segment threshold
design.CTS.50T <- make.design(region = global.CTS,
                  transect.type = "line",
                  design = "segmentedgrid",
                  design.angle = 90,
                  spacing = c(10500),
                  seg.length = 7500,
```

```
                seg.threshold = 50,
                edge.protocol = "minus",
                truncation = 150,
                coverage.grid = cover.CTS)

# Generate example transects
eg.transects <- generate.transects(design.CTS.50T)
# Plot the transects with the stratified region
plot(region.CTS, eg.transects)
```

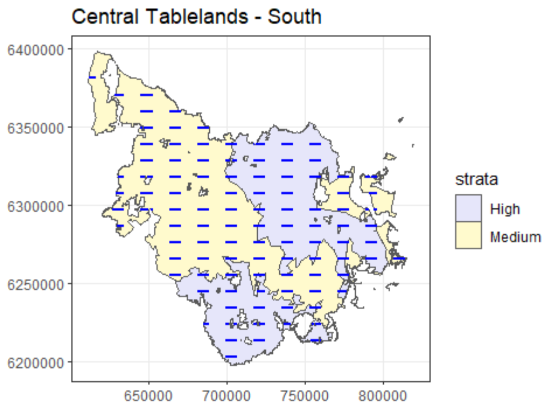


Figure A9. Example transects from a design which generates transects using a global region using a segment threshold of 50.

The coverage scores are now uniform across most of the management zone and the edge effects around the exterior of the management zone are much less noticeable (Figure A10).

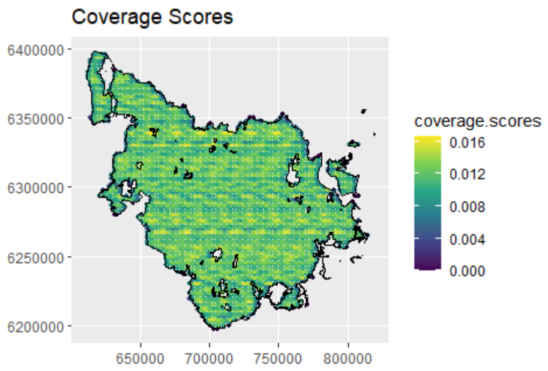


Figure A10. Coverage scores across the Central Tablelands south management zone for the global design using a threshold of 50.

The histogram of the coverage scores for the 50% threshold design looks reassuring; although there is still a left skew, the majority of coverage scores look to have similar values of around the median value of 0.012 (Figure A11). For this design only 0.4% of coverage grid points have coverage scores of 0, 1.4% have coverage scores of less than 0.005 and 10.3% have coverage scores less than 0.009.

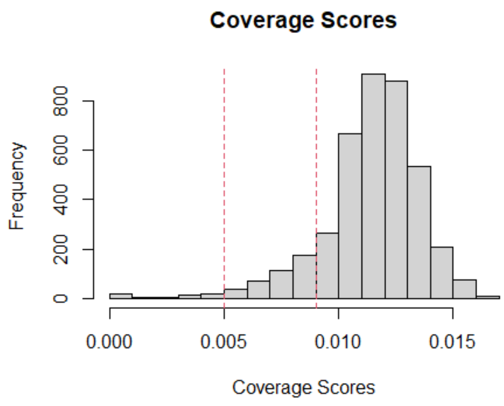


Figure A11. Histogram of the coverage scores across the Central Tablelands south management zone for the global design using a threshold of 50. The arbitrarily chosen values of 0.005 and 0.009 are indicated by red dashed lines.

Discussion

Modifying the design has achieved more uniform coverage scores across the study area. This will have greatly reduced the potential for over or under estimation of animal abundance due to animal density differing in the areas of lower coverage, as these areas now make up a much smaller proportion of the study region.

It is likely that the final design will have a higher effort and cost than the original, especially with the need to fly incomplete samplers around the perimeter of the management zone. To account for this the design spacing may need to be increased. One suggestion to improve efficiency which should also further improve coverage scores would be to fly the remainder of the incomplete transects a fixed distance along the study region boundary. An example of this is shown in Figure A12. The practicalities of doing this will need to be assessed in terms of additional cost versus benefit trade-off for this study. The additional lengths of transect need to be located such that they are far enough away from the area which has just been surveyed so as not to have been disturbed by the passing of the helicopter.

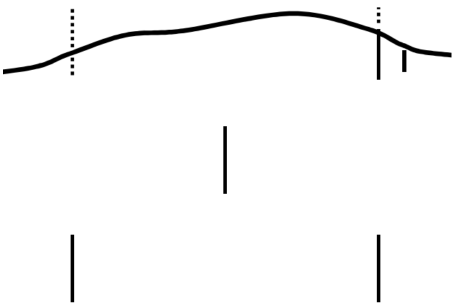


Figure A12. Samplers less than half inside the study area are discarded, those more than half inside are surveyed and the additional sampler length is surveyed at some distance along the boundary after being reflected back inside.

The final consideration of the suggested design is the independence of samplers. Samplers that lie across stratum boundaries will be divided into two separate samplers which will then be treated as independent. This may result in a small underestimation of the variability. In practice, we expect the effects to be negligible due to the fact that the majority of samplers will fall within the central parts of each stratum, with only a few crossing internal stratum boundaries.

Western Plains

General design review

The Western Plains design involved randomly laying down a grid of 50 km by 20 km rectangles. A subset of these rectangles was randomly chosen, and an equal spaced zigzag design was laid within each of the selected rectangles (Figure A13). The main concern of the Western Plains survey design is the reliability of the variance estimates due to the low numbers of plots sampled in each management zone; between four and seven. Although there are multiple transects laid down within each sampled block, the encounter rates on these are likely to be more similar than those in a block in a different part of the management zone. For the purposes of analysis, it is therefore more appropriate to think of each block as a sampling unit rather than each individual transect within a block. Buckland *et al*. (2001) recommended that for reliable estimation of variability a minimum of 10-20 independent samplers per stratum.

Within the blocks, a fixed zigzag pattern was used but this would not impact the coverage due to the randomisation of the grid location. In addition, an equal spaced zigzag design gives uniform coverage within a rectangular region. This design would therefore give close to uniform coverage across the whole study region, with the exception of a narrow band of lower coverage around the stratum and study region boundaries due to standard line transect edge effects. As discussed above, due to the ratio of study region boundary to area we expect any effects on estimates to be negligible.

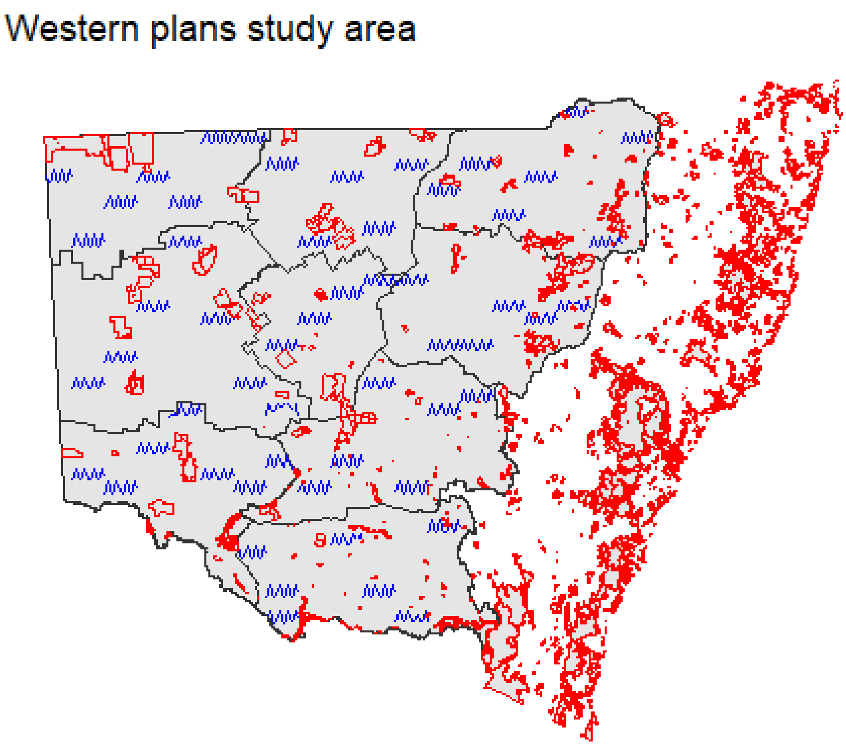


Figure A13. The Western Plains management zones (defined in grey), national parks (defined in red) and transects (blue lines).

The factors which result in non-uniform coverage for this design was the exclusion of national parks and the inability to obtain permission to fly over some privately owned land. The survey was planned to avoid national parks. Any selected grid cell which was more than 50% inside a national park was excluded and any smaller area of national park which lay within a selected block was also avoided. Given the difficulties in automating the survey design, it is not possible to run explicit coverage simulations for this region, however, these additional factors will have led to lower areas of coverage around the boundary of national parks and no coverage within national parks themselves, or areas where permission to fly could not be obtained.

Omitting all national parks but then later including the areas of these in the analyses could cause over or under estimation of animal abundance if animal density differs in these regions compared to the rest of the study area. Given that no data is currently collected within national parks it is unknown as to whether animal density may vary in these regions. The areas excluded due to a lack of landowner permission are less problematic. It could be reasonable to assume these are missing at random and should therefore be unlikely to affect the accuracy of the estimates.

In addition to these areas of lower coverage, there are also small areas which will have higher than average coverage. When part of a selected rectangular plot cannot be surveyed in full, the zigzags are amended to exclude that area. This often involves several shallower angled transects being used which will result in a region of higher coverage within the sampled rectangle (Figure A14).

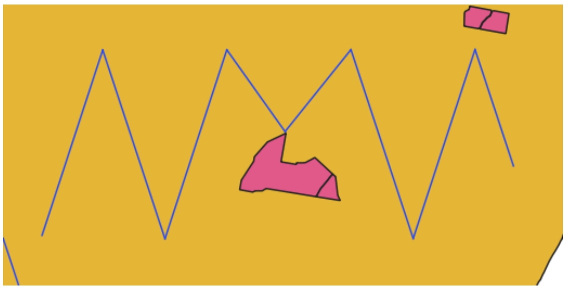


Figure A14. Non-uniform coverage zigzags within a rectangle.

A preferable approach would be to simply clip the transects from the regions where surveying cannot be undertaken (Figure A15). This will maintain uniform coverage across the surveyed rectangles. In the previous approach, while the area of higher coverage may seem small in relation to the study area, in some management zones it can make up a significant proportion of the covered area as the number of sampled rectangles is only between 4 and 7. If animal density differs in these regions of higher coverage, it could cause the abundance to either be over or underestimated. This is more likely to have a substantial impact if kangaroo density is highly aggregated within the survey region and likely to be less important if trends in animal density are smoother through space. Also with this alternative approach, if permissions allow, data could be collected along the full length of the transect lines, including within the national parks, but later be omitted from the main analyses. This would allow abundance in the national parks to be assessed separately.

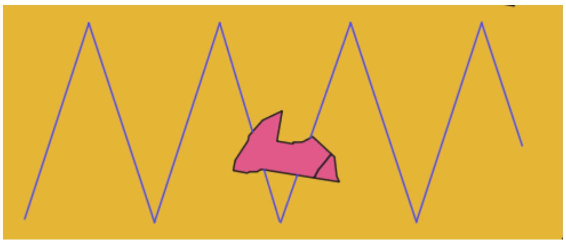


Figure A15. Uniform coverage zigzags within a rectangle.

Spatial Modelling

A spatial model fitted to the 2021 Western Plains survey data was provided by the KMP as a basis for describing animal distribution throughout the study area to help assess the adequacy of the survey design. The spatial model used additional spatial covariates related to habitat to try and more accurately model kangaroo distribution within the Western Plains region. This model was going to be used to assess surveys using drones and was therefore fitted to the counts of both the red and grey kangaroos combined. It was noted that the estimate of the combined abundance of red and grey kangaroos across the whole of the Western Plains was lower than the estimate based on the mark recapture distance sampling (MRDS) model alone. It also appeared that the confidence intervals for the two methods did not overlap, however, it is likely that the variability of the MRDS estimates was under-estimated due to the assignment of transect ID as sampling units rather than blocks. We used this model to check whether the proportions of the different habitat types included in the survey matched those present throughout the Western Plains region.

Assessment of candidate spatial covariates

The following covariates were provided for consideration in the density surface model (dsm): Elevation, Slope, NDVI, Cover Class and Vegetation Form. The values of these covariates have been investigated within the region covered by the survey (the segment data) and the prediction grid.

Elevation: The range of elevation values in the segment data is between 25m and 637m. Elevation values in the prediction grid range from 21m to 2,127m with 0.86% of the prediction covariate values lying out with the range of the segment data.

Slope: The range of the slope values in the segment data is between 0 and 40. Slope values in the prediction grid range from 0 to 100 with 0.17% of the prediction covariate values lying out with the range of the segment data.

NDVI: The covariate values in the Nov21_NDVI column were assumed to be the correct values for these analyses. In the segment data, the NDVI values ranged from -0.16 to 0.82, while in the prediction grid the range of values was from -0.20 to 0.91. Only 0.10% of the prediction values were out with the range of the segment data values.

Table A1. Cover Class: The percentage of segment data points and the percentage of prediction data points which fall into each cover class category.

Cover Class	A	B	C	D	E	F	G
Segments	29.2	10.7	12.0	23.7	21.1	2.4	0.9
Prediction	27.1	9.7	11.9	24.7	21.9	3.1	1.6

Table A2. Vegetation Form: The percentage of segment data points and the percentage of prediction data points which fall into each vegetation form category.

Vegetation Form	Segments	Prediction
Arid Shrublands (Acacia sub-formation)	15.9	14.7
Arid Shrublands (Chenopod sub-formation)	12.0	11.9
Dry Sclerophyll Forests (Shrub/grass sub-formation)	0.2	1.0
Dry Sclerophyll Forests (Shrubby sub-formation)	2.2	2.1
Forested Wetlands	0.8	1.2
Freshwater Wetlands	4.1	3.9

Vegetation Form	Segments	Prediction
Grasslands	4.6	4.2
Grassy Woodlands	1.9	2.4
Not native vegetation	27.1	29.2
Saline Wetlands	2.0	1.7
Semi-arid Woodlands (Grassy sub-formation)	6.0	7.7
Semi-arid Woodlands (Shrubby sub-formation)	21.1	21.9
Wet Sclerophyll Forests (Grassy sub-formation)	0.0	0.3

Discussion

Tables A1 and A2 suggest that the current survey design is producing a representative sample of the study area in terms of vegetation. This can be seen in the fact that the percentages of each category for both cover class and vegetation form are similar across the segment and prediction datasets. It should be noted that this has only been assessed across the region as a whole and not individually at the KMZ level.

The ranges of the values for NDVI largely overlap between the segment data and the prediction data. There are however some higher values for elevation and slope within the prediction data which will involve significant extrapolation beyond the range of the covariate values in the segment data. Predictions in these regions will not be supported by any data. The combinations of covariates have not been considered here but can be assessed using the R package dsmextra (Bouchet *et al*., 2020).

We also note that the modelling for the 2021 red kangaroo data in Appendix B has resulted in lower estimates of abundance. These lower estimates would mean that the confidence intervals for the total number of kangaroos (both red and grey) across the whole region would now overlap to some extent. In addition, the confidence intervals for the MRDS estimates will be larger with the revised definition of sampling units, further increasing the overlap of the intervals.

Appendix B: Model selection

This appendix investigates the effect of selecting from a wider pool of candidate models on abundance estimates. Survey data on red kangaroos collected in 2021 during the Western Plains survey (Curtis and McLeod, 2021) are used.

Mark-recapture distance sampling (MRDS) methods were used on this survey and, since point independence was assumed, the analysis consisted of fitting a model for the distance sampling (DS) component and one for the mark-recapture (MR) component. The MR model provides the intercept (i.e. $g(0)$) for the DS model but Akaike information criterion (AIC) values are calculated independently for each model. AIC provides an objective method for choosing between models.

Curtis and McLeod (2021) fitted 17 combinations of models to the data and selected between them based on AIC. These combinations should provide a variety of possible options but since the two models (DS and MR) are fitted independently in the analysis, the actual number of models tried for each component was smaller, for example, there were 4 models for the DS component and 12 for the MR model. In this appendix, a more extensive search for an optimal model is described and the results obtained from the two approaches compared. It should be noted that the data used here was slightly different (a few incomplete records were excluded (pers. comm.)) to that used by Curtis and McLeod (2021) and so to make the results comparable, the model selected by Curtis and McLeod (2021) has been refitted to the data and updated abundance estimates obtained. The models selected by Curtis and McLeod (2021) have been used as a comparison, but these exclusions may have resulted in a different model choice by these authors.

Methods

There are many approaches to model selection, for example forward or backward selection and all-possible-subsets and they are not guaranteed to result in the same model choice. An alternative method is to carefully select a set of models and use an objective criterion to choose between them. In this analysis, forward selection was used for two reasons; a lack of *a priori* knowledge of which explanatory variables (or combination of variables) would be useful/biologically plausible and, with so many detections (see Table B1a), fitting models with a lot of explanatory variables (as in backward selection) was time consuming.

Forward selection begins with a 'null' or intercept only model. Each of the potential variables were included in turn and the variable which provided the largest improvement in the model was selected. Starting with this new model, each of the remaining variables were then included in turn and so on until there was no improvement in the model. Here, AIC was used to quantify the models and the model with the lowest AIC was selected, noting that if a simpler model (i.e., one with fewer parameters) was within 2 AIC units of a more complicated model, the simpler model was selected. Interaction terms can also be considered, given a sufficient sample size, but only a limited number of interaction terms were included here and so there may be a more optimal model. Also, other variables may be useful to consider such as taking the logarithm of size, particularly if there are a few very large groups (mobs). The aim here was to use an alternative model selection procedure and compare results rather than obtain the optimal model.

Forward model selection was used to select the DS and MR models independently. Both the half-normal (HN) and hazard rate (HR) forms of the detection function were used (again assuming no *a priori* knowledge of which might be more appropriate). However, results using both the HN and HR are provided to assess the difference this might have on abundance estimates.

In addition to perpendicular distance, the candidate explanatory variables were:

- IBRA – bio region (10 levels)
- K.Class – vegetation type derived from Keith classes (6 levels)
- size – group size (from 1 to 53 animals)
- observer – position of observer, used in the MR model only (1 or 2)
- obsname – individual observers (10 levels)
- session – time of day (am or pm)
- glare – presence of glare (0 or 1)
- bearing – bearing of aircraft (from 15° to 345°)

There were two variables measuring vegetation, IBRA and K.Class, and so because there was some similarity between them (pers. comm.), when one was included in the model, the other was then excluded from further consideration.

Once variables had been selected, interaction terms were considered in the MR model.

Survey data

These data were described and analysed by Curtis and McLeod (2021); a summary of the survey data is shown in Table B1. Nearly 8,555 km of search effort was conducted and a total of 3,123 groups detected. Of these, 517 groups were detected by both observers (Table B1b). Note that the zigzag transects were used as the sampling unit in the variance estimate of the encounter rate. The choice of sampling unit is discussed elsewhere in this review.

A group of two animals was the group size most frequently detected (38%, Table B1c) and 70% of all groups were either 1, 2 or 3 animals but groups of up to 53 animals were detected.

Table B1. Summary of 2021 survey data.

a) Area of KMZ, number of individual zigzag transects (k), search effort (L), search region covered during the survey (i.e. $2wL$ where w is the truncation distance of 300 m), number of detected groups of red kangaroos (n), mean group size ($\bar{s}$) and maximum group size (s) in each kangaroo management zone (KMZ).

KMZ	Area (km^2)	k	L (km)	Covered region (km^2)	n	$\bar{s}$	Maximum s
Bourke	55,019	35	692.3	415.4	117	2.19	8
Broken Hill	91,044	57	1,111.7	667.0	912	3.24	28
Cobar	40,419	32	637.3	382.4	158	2.82	12
Coonabarabran	61,711	48	932.6	559.6	101	5.65	52
Griffith North	65,758	48	959.1	575.5	230	4.00	32
Griffith South	64,126	56	1,068.7	641.2	134	3.42	16
Lower Darling	56,599	55	1,068.0	640.8	959	3.59	53
Narrabri	65,755	57	1,070.1	642.4	227	3.78	32
Tibooburra	54,849	55	1,014.4	608.6	285	2.26	8
Total	555,280	443	8,554.7	5,132.8	3,123	3.38	53

b) Number of groups detected by front (1) and rear (2) observers, and the number detected by by both observers (duplicates).

Seen by	Number of groups
Observer 1	1,867
Observer 2	1,773
Both 1 and 2	517

c) Frequency of detected group sizes. The maximum number in a detected group was 53.

Group size	Frequency
1	488
2	1,196
3	543
4	359
5	145
6-9	265
10-19	100
20-20	17
30-39	7
40-40	1
50-59	2
Total	3,123

Results

The AIC values for the two component models are provided at the end of this appendix in Table B4. The variables selected in the final model (which is referred to as Model 1) were:

- DS model (HR form): IBRA + size + obsname + session
- MR model: obsname + size + distance + IBRA + obsname:distance

The total AIC for Model 1 was 14,241.38. Purely on the basis of AIC, a HR detection function form was selected and, for comparison, a model with the same variables as Model 1 but using a HN form was also fitted (referred to as Model 2).

Model 1 contained similar variables to that selected by Curtis and McLeod (2021) – these authors selected the following model (referred to as Model 3):

- DS model (HN form): obsname + K.Class
- MR model: obsname + distance + observer + IBRA + obsname:distance + observer:distance

The total AIC for Model 3 was 14,338.85. Note that this value was slightly different to that reported by Curtis and McLeod (2021) because of the slight differences in data – these differences are reflected in the results and plots shown here obtained with Model 3.

With so many parameters, goodness of fit was assessed visually (Figure B1). Ideally, points should fall along the straight line and Model 1 suggests a slight improvement over Model 3.

Table B2a shows that, although there were some differences in the variables included in the MR model, the estimate of detection probability on the transect for at least one observer (which is used to scale the DS detection function model) was the same for Model 1 and Model 3.

The average probability of detection was similar for Models 2 and 3 which both used a HN detection function form. It was higher for Model 1 which used a HR detection function form (Table B2b).

Table B2. Estimated probabilities of detection from the MR and DS component models.

a) Estimated probability of detection on the transect for each observer and by at least one observer (i.e., for observer 1 or 2 or both) from the MR models in Model 1 and Model 3. CVs are provided in parentheses. The estimates for Model 2 are the same as Model 1.

Observer	Model 1	Model 3
1	0.34 (0.06)	0.31 (0.07)
2	0.34 (0.06)	0.37 (0.07)
Both 1 and 2	0.55 (0.05)	0.55 (0.05)

b) Average probability of detection (over all distances and other variables) from the DS model unadjusted for detection on the transect (i.e., assuming detection on the transect was certain) and adjusted for uncertain detection (using estimates from Table B2a).

Correction	Model 1	Model 2	Model 3
Unadjusted	0.62 (0.02)	0.54 (0.02)	0.55 (0.02)
Adjusted	0.34 (0.06)	0.30 (0.06)	0.30 (0.06)

The higher probability of detection with Model 1 has resulted in lower abundance estimates (Table B3) compared to Models 2 and 3. Model 3 has resulted in higher abundance estimates compared to Model 2 in all KMZ apart from Bourke.

Table B3. Abundance estimates for red kangaroos in 2021 obtained with Models 1, 2 and 3: coefficients of variation (CV), and lower (LCL) and upper limits (UCL) of the 95% confidence interval.

a) Model 1

KMZ	Abundance	CV	LCL	UCL
Bourke	159,005	0.21	106,107	228,277
Broken Hill	1,042,714	0.10	852,528	1,275,327
Cobar	156,201	0.23	99,515	245,179
Coonabarabran	143,724	0.25	88,115	234,427
Griffith North	208,131	0.28	119,916	361,543
Griffith South	81,522	0.34	41,491	160,173
Lower Darling	574,774	0.12	456,203	724,162
Narrabri	222,544	0.18	157,472	314,506
Tibooburra	270,672	0.18	189,708	386,189
Total	2,859,287	0.07	2,491,811	3,280,956

b) Model 2

KMZ	Abundance	CV	LCL	UCL
Bourke	178,979	0.20	119,609	267,817
Broken Hill	1,201,010	0.10	981,588	1,469,480
Cobar	180,726	0.23	114,855	284,275
Coonabarabran	162,184	0.25	99,997	263,044

Griffith North	234,249	0.28	134,836	406,958
Griffith South	91,304	0.35	46,437	179,521
Lower Darling	645,517	0.12	511,792	814,182
Narrabri	255,975	0.18	180,240	363,532
Tibooburra	309,171	0.18	215,794	442,952
Total	3,259,113	0.07	2,836,270	3,744,996

c) Model 3 (after Curtis and McLeod, 2021)

KMZ	Abundance	CV	LCL	UCL
Bourke	163,527	0.21	108,493	246,478
Broken Hill	1,279,375	0.11	1,023,616	1,599,038
Cobar	178,806	0.23	114,284	279,756
Coonabarabran	224,114	0.32	120,378	417,243
Griffith North	284,108	0.30	158,522	509,185
Griffith South	113,595	0.34	58,735	218,698
Lower Darling	749,656	0.12	587,012	957,363
Narrabri	306,786	0.19	209,963	448,260
Tibooburra	335,528	0.19	233,594	481,942
Total	3,635,486	0.08	3,132,496	4,219,265

Discussion

The key information required from the MR model is the probability of detection on the transect by at least one observer and in this example, this estimate was the same with both Model 1 and 3 even though the variables selected were slightly different (Table B1). In the DS model, more variables were selected in Model 1 than in Model 3, and a hazard rate detection function form was selected over the half-normal form.

Forward selection was used to select Model 1 and this resulted in the DS and MR component models containing several explanatory variables – backwards selection could also be implemented to test whether all variables were still important in the model fit, but this was not done here. This rather naïve approach to which variables (or combination of variables) would be important, resulted in a large number of models being fitted. It is likely that a more informed approach would result in fewer models being fitted, nevertheless selecting from a wider pool of component models may result in a more optimal model.

A hazard rate detection function form was not considered by Curtis and McLeod (2021) but, on the basis of AIC alone, a hazard rate was shown to provide a better fit to the pooled (i.e., over all observers) detection function for these data. However, the KMP favour the half normal over a hazard rate form given practical experience of survey techniques and observer behaviour:

- A half normal form better reflects the way in which observers should search, and this is the case for experienced observers (for example Fig. B4a).
- Some less experienced observers were recording a similar number of detections, or more detections, in the second distance bin compared to the first bin (for example Fig. B4b). This may be because they were not spending enough time looking in the first bin or were slow to record the location (because they wanted to be sure of the species) and so animals had moved by the time observers were recording the data. A half normal form adjusts for this by smoothing through high numbers of detections in the second bin.

The result of selecting a HN detection function with these data has resulted in slightly higher abundance estimates.

The coefficients of variation (CVs) were very similar with each model and there was substantial overlap in the confidence intervals (CI) between the estimates from all models (Figure B3), bearing in mind that the CIs may be underestimated (this is discussed elsewhere in the report). Of course, the true abundances are unknown, however, given the large number of detections, small differences in the estimated detection probability can result in substantial absolute differences. In this example, using the estimates from Model 1 would be the precautionary approach.

This model selection has been performed on detections of red kangaroos – grey kangaroos have not been considered but the KMP confirmed that similar issues also occur with grey kangaroos.

Recommendations

1. Record the distance bin first rather than species so that animals do not have time to respond before the bin gets recorded.
2. Record information on movement of the detected group so that responsive movement can be investigated. This may be an easier option to implement than recommendation 1 and inform whether recommendation 1 would be worth implementing.
3. Investigate a wider pool of candidate models.

Figure B1. Goodness of fit for Model 1 and Model 3.

a) Model 1 b) Model 3

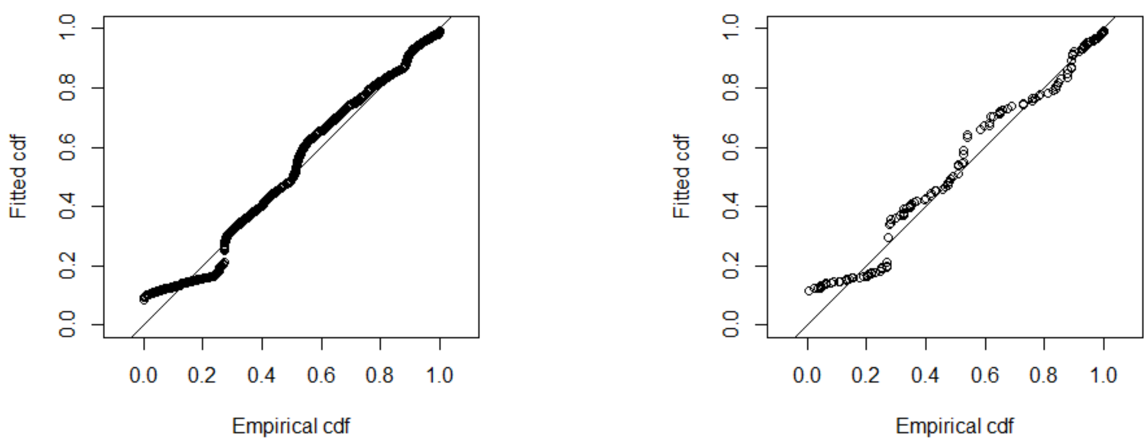
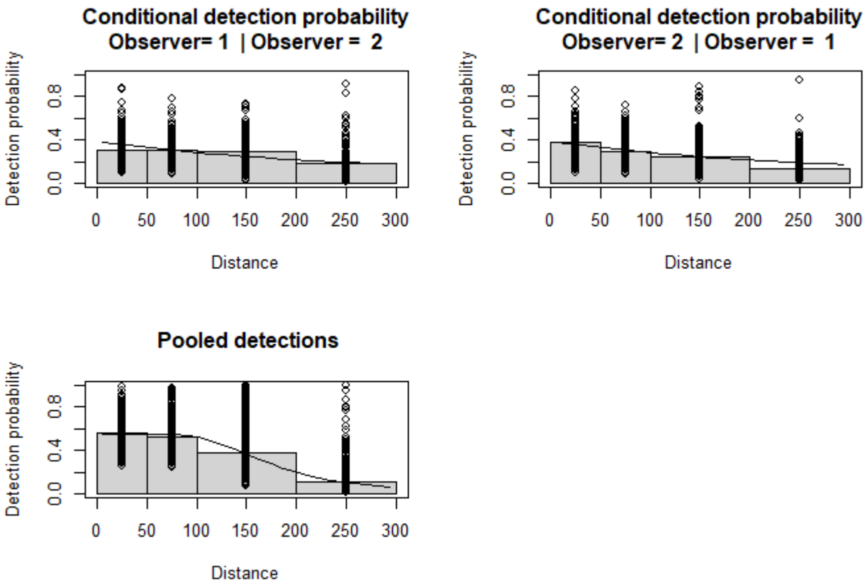


Figure B2. Plot of fitted MR model (top) and DS model (bottom) for a) Model 1 and b) Model 3. Circles indicate individual detections and the smooth lines indicate the average fitted detection functions.

a) Model 1



b) Model 3

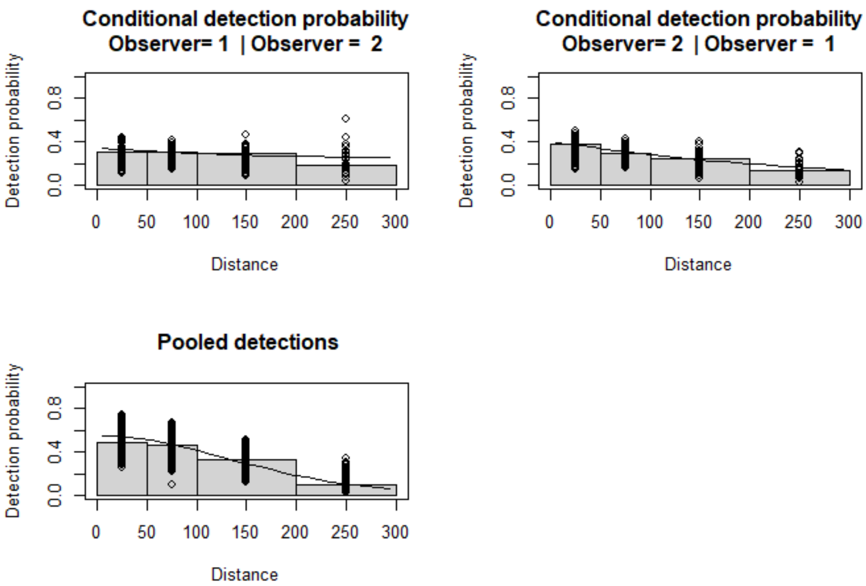


Figure B3. Comparison of red kangaroo abundance estimates for Models 1, 2 and 3. Vertical lines indicate 95% confidence intervals.

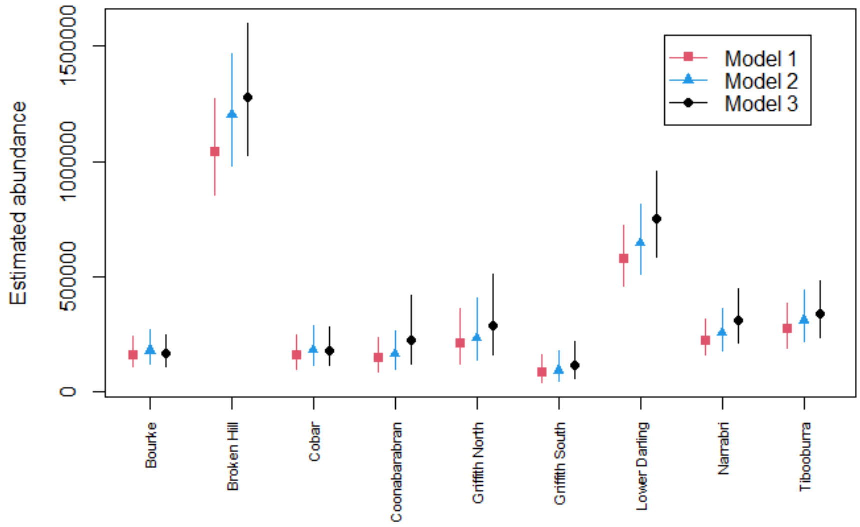


Figure B4. Examples of detection functions of individual observers. A half normal detection function has been fitted and the intercept assumes that detection at zero distance is certain.

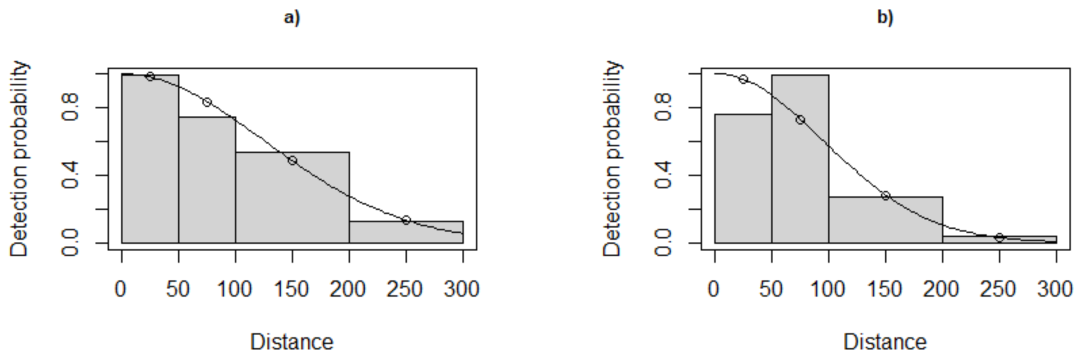


Table B4. AIC values obtained for the a) MR and b) DS component models.

a) MR model

Covariates included	**AIC**
None	6417.60
One variable	
obsname	6356.02
size	6380.86
distance	6394.27
IBRA	6403.02
observer	6416.20
session	6419.36
K.Class	6419.49
bearing	6419.55
glare	6419.59
Two variables	
obsname + size	6316.48
obsname + distance	6330.99
obsname + IBRA	6341.72
obsname + observer	6355.79
obsname + glare	6357.70
obsname + session	6357.83
obsname + bearing	6357.87
obsname + K.Class	6358.53
Three variables	
obsname + size + distance	6280.37
obsname + size + IBRA	6303.91
obsname + size + observer	6316.30
obsname + size + glare	6318.20
obsname + size + session	6318.36
obsname + size + bearing	6318.45
obsname + size + K.Class	6320.03
Four variables	
obsname + size + distance + IBRA	6264.48
obsname + size + distance + observer	6279.98
obsname + size + distance + K.Class	6281.27
obsname + size + distance + glare	6282.15
obsname + size + distance + session	6282.18
obsname + size + distance + bearing	6282.26
Five variables	
obsname + size + distance + IBRA + observer	6263.93
obsname + size + distance + IBRA + session	6266.41
obsname + size + distance + IBRA + bearing	6266.14
obsname + size + distance + IBRA + glare	6266.26
Interactions	
obsname + size + distance + IBRA + obsname:distance	**6177.77**
obsname + distance + observer + IBRA + obsname:distance + observer:distance[1]	6196.59

[1] Model selected by Curtis and McLeod (2021)

b) DS model. AIC values are provided for both half normal (HN) and hazard rate (HR) forms of the detection function. Fitting the variable bearing created warning messages and so was not considered further.

Covariates included	**HN**	**HR**
Distance only	8197.80	8179.63
One variable (in addition to distance)		
IBRA	8125.27	8104.12
K.Class	8164.99	8144.13
obsname	8168.65	8144.13
size	8166.99	8144.30
session	8197.83	8179.42
glare	8198.70	8180.98
bearing	problems	
Two variables		
IBRA + size	8109.94	8091.10
IBRA + obsname	8115.44	8092.00
IBRA + log(size)	8116.59	8092.63
IBRA + session	8125.15	8104.22
IBRA + glare	8127.02	8106.00
obsname + K.Class	8142.83[1]	8117.10
Three variables		
IBRA + size + obsname	8094.87	8066.46
IBRA + size + session	8110.21	8086.54
IBRA + size + glare	8111.75	8087.88
Four variables		
IBRA + size + obsname + session	8092.31	**8063.62**
IBRA + size + obsname + glare	8095.93	8067.93
Five variables		
IBRA + size + obsname + session + glare	8093.39	8065.17

[1] Model selected by Curtis and McLeod (2021)

Appendix C: Adjustment factors

Various factors are used as part of the analyses to determine 1) western and eastern grey populations from estimates of grey kangaroos and 2) adjust common wallaroo estimates. This appendix describes how these factors were obtained and illustrates their use to ensure that the methods being applied are fully understood. Some comments and recommendations for consideration are also provided and these are also listed in the main report.

Ratio of eastern and western grey kangaroos

Both eastern grey kangaroos (EGK) and western grey kangaroos (WGK) are found in the Western Plains but they cannot be determined to species during the surveys. Abundance estimates are obtained for grey kangaroos for each KMZ and then these estimates are divided into an estimate for eastern and western greys based on ratios of eastern to western greys determined from spotlight surveys conducted in each KMZ (Cairns and Gilroy, 2001). The proportions used to split the 2022 survey abundance estimates are shown in Table C1 (derived from estimates provided in Curtis and McLeod, 2022).

Table C1. Abundances of eastern and western grey kangaroos in each KMZ (from Curtis and McLeod, 2022) and the percentages used to split the abundance estimate of grey kangaroos.

KMZ	Area (km^2)	Total grey kangaroos	EGK	WGK	EGK:WGK percentage
Bourke	55,019	30,385	18,839	11,546	62:38
Broken Hill	91,044	211,832	88,970	122,863	42:58
Cobar	40,419	40,278	16,917	23,361	42:58
Coonabarabran	61,711	1,042,877	969,876	73,001	93:7
Griffith North	65,758	941,071	781,089	159,982	83:17
Griffith South	64,126	535,660	444,598	91,062	83:17
Lower Darling	56,599	340,884	105,674	235,210	31:69
Narrabri	65,755	289,798	286,900	0	99:0
Tibooburra	54,849	16,510	13,208	3,302	80:20
Total	555,280	3,449,296	2,726,071	720,327	

Comments/Recommendations

- An estimate of WGK was not included for Narrabri and so the sum of EGK and WGK did not match the total for grey kangaroos.
- The uncertainty associated with the proportions is not incorporated into the uncertainty associated with the abundance estimate and so total uncertainty will be underestimated. The CV of the proportion can be incorporated into CV of the abundance estimate using the delta method (Seber 1982) assuming that these two components are independent. This seems reasonable since they have been derived from two different types of survey.
- A CV for the estimated proportion ($\hat{p}$) can be found from the following equations:

  $se(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ and $cv(\hat{p}) = \frac{se(\hat{p})}{\hat{p}}$

  where n is the number of observations used to obtain the estimated proportion $\hat{p}$.

- The proportions obtained from the survey in 2000 had changed compared to a previous survey undertaken 16 years earlier in 1984 (Caughley *et al*. 1984) (Cairns and Gilroy, 2001).

Given that 20 years have now passed since the survey on which the current ratios are based, it is advisable to check whether these proportions are still valid.

- The proportions are obtained from spotlight surveys and then applied to the aerial survey data. This implies that the proportion of EGK to WGK detected in the aerial survey, is the same as in the spotlight surveys. Is this a reasonable assumption?
    - The survey in 2001 involved driving along roads to spot kangaroos (which may be the only option if surveys are undertaken in poor light when kangaroos are most active). However, if one species prefer habitat closer to roads the sample would contain a higher proportion of one species than over the whole region of interest.
- In general, any survey would also need to follow good survey design practise, such as randomly choosing half-degree blocks. It was not clear how the blocks used in 2000 were chosen.

Correction multiplier of common wallaroo estimates

Conventional distance sampling estimates from the helicopter surveys in the Tablelands may underestimate the population because observers are not detecting everything on the transect and because some animals may not be available for detection. Thus, the resulting estimates will underestimate the true value. To account for this, Cairns *et al*. (2020) adjusted wallaroo population estimates obtained from the helicopter surveys by multiplying by 1.85 (=1/0.54). Based on a study in Queensland, Clancy *et al*. (1997) indicated that estimates from helicopter line transect surveys were lower than estimates obtained from walked line transect surveys: the relationship between density estimates obtained from the helicopter ($\widehat{D}_h$) and walked line transects ($\widehat{D}_w$) was $\widehat{D}_h = 0.54\widehat{D}_w$. Cairns *et al*. (unpublished) found similar estimates from a survey in NSW.

Comments/Considerations

- This factor is applied without incorporating uncertainty of the estimated parameter into the uncertainty of the abundance estimate. The estimated parameter in the regression equation will have an associated standard error term which could be used to obtain a CV, which in turn could be incorporated into the CV of the abundance estimate using the delta method (Seber, 1982).
- One question relating to the study was whether the terrain in these comparative studies was similar to the terrain in the regions where it is now being applied? For example, if the comparative study was conducted in a region where there was substantial tree cover, then it might be expected that the number seen during a walking survey would be higher than the number seen during the aerial survey. If the helicopter surveys are now undertaken in regions where there is little cover, the observers may be detecting a higher proportion than would be anticipated if there was more cover.
- The adjustment factor was based on a survey over 20 years ago and the difference between walking and helicopter surveys may have changed over the years (e.g., more/less habitat cover).
- Why is adjustment made for wallaroos but not for the other species identified during these surveys?
- The walking surveys are being used to obtain the true number of kangaroos, but could this type of survey be subject to issues such as responsive movement or wallaroos moving ahead of the observers and so being counted twice?
- One possible solution would be to attach GPS tags to a sample of wallaroos and then (knowing where they are because of the GPS tag) fly directly (at zero distance for observers)

over them to determine what proportion of tagged animals were detected by the observers. This could then be used as an adjustment factor to correct for both perception and availability. The helicopter would also need to fly over transects where there were no tagged wallaroos to avoid cueing observers. This would be costly and so likely benefits would need to be considered.

- An adjustment factor is not applied to estimates of EGK, however, unless the probability of detection is certain (or very close to one) on the transect, these estimates will underestimate the true abundance. The above solution may also be used to determine the proportion detected of EGK.

Appendix D: Variance estimation

The variance, and the resulting coefficient of variation (CV) and confidence interval (CI), associated with an abundance estimate measure the uncertainty associated with the estimate. The variance of the encounter rate can frequently contribute most to the overall variance of an abundance estimate. One component of the encounter rate (ER) variance is the number of independent sampling units, or samplers and with a small number of samplers the variance may not be estimated reliably. This appendix discusses variance estimation for the Tablelands and the Western Plains surveys.

Tablelands surveys

With systematic survey designs (i.e., equally spaced samplers with a random start point), the samplers are frequently treated as independent for the purposes of calculating the variance. Fewster *et al*. (2009) showed that using this approach, where there is a spatial trend in a population, can result in overestimating the variance. They provided alternative estimators to reduce this overestimation which used poststratification (i.e., grouping adjacent samplers).

In the Tablelands, a segmented grid design is used (where samplers are short and laid on a systematic grid) and so these methods may be of interest, although it is not clear how poststratification would be applied with this design. However, Fewster *et al*. (2009) derived an estimator for point transects which they stated could also be used where lines are short, of equal length and positioned on a regular grid (Fewster *et al*., 2009, Appendix B). This approach may reduce the variance but would require programming because, currently, this variance estimator is not available in Distance for Windows (Thomas *et al*., 2010) or in the distance sampling R packages. It may also require some adaptation for unequal length lines (which might occur if samplers are split where they cross a boundary).

Western Plains surveys

In the Western Plains surveys, 56 blocks throughout 9 KMZ are sampled. Within each block there are 8-9 zigzag transects. In analysis of these surveys (e.g., Curtis and McLeod 2021, 2022), each zigzag transect was used as an independent sampling unit to calculate variance of the encounter rate. Within blocks, these transects are not located independently of each other, however blocks were selected at random, and so using blocks as the sampling unit has been suggested. Because there are typically fewer blocks per stratum than the number of sampling units recommended for reliable variance estimation (around 20 (Buckland *et al*., 2001)), this may produce unreliable estimates of encounter rate variance, and hence unreliable estimates of the variance of density estimates. The effect of this change can be seen in Table D1.

This appendix describes a method for estimating the variance with a model-based approach that uses information from all strata to estimate the encounter rate variance in each stratum. This increases the number of sampling units available to estimate the variance in each stratum, but to do this it introduces an assumption that the ratio of the encounter rate variance to the encounter rate mean is the same in all strata (or KMZ in this application). We implement this method and compare estimates to those from the design-based approach. This assumption may not be valid across all KMZ, if for example, the habitat is very heterogeneous, and so the KMZ are divided into two groups such that the assumption may be more likely to be valid.

The design-based, encounter rate variance of the encounter rate for each KMZ is given by:

$$\widehat{var}(n_i) = \frac{L_i \sum l_{ij} \left({n_{ij}}/{l_{ij}} - {n_i}/{L_i} \right)}{k_i - 1}$$

where for each KMZ i,

- n_i is the total number of groups,
- L_i is the total search effort,
- k_i is the number of blocks,
- n_{ij} is the number of groups in block j, and
- l_{ij} is the search effort in block j.

The coefficient of variation (CV) of the number of groups is given by:

$$\widehat{CV}(n_i) = \frac{\sqrt{\widehat{var}(n_i)}}{n_i}$$

Note, the CV of the encounter rate (as reported in summary output from the R function dht; Laake *et al*., 2022) will be the same as the CV of the number of groups.

An empirical estimate of the ratio of the variance to the expected number of groups is given by:

$$b_i = \frac{\widehat{var}(n_i)}{n_i}$$

Data

Table D1 shows that number of blocks within each KMZ and the number of zigzag transects. Within each KMZ, the number of blocks is less than the recommended number of 10-20 (Buckland *et al*., 2001). The corresponding increase in encounter rate CV is shown in Table D1.

Table D1. Summary of 2022 survey data and design-based estimates: number of individual zigzag transects (k), search effort (L), number of unique groups (n) and the CV of the encounter rate using transects and blocks as the independent sampler. Group indicates the division of KMZ into two types such that the ratio of ER variance to ER mean is the same.

KMZ	Group	Area (km^2)	Number of blocks	k	L (km)	n (groups)	n/L	Transects $CV({n}/{L})$	Blocks $CV({n}/{L})$
Bourke	1	55,019	5	41	802.7	204	0.254	0.14	0.25
Broken Hill	1	91,044	7	57	1,086.6	1,485	1.367	0.07	0.13
Cobar	1	40,419	4	31	614.5	172	0.280	0.21	0.41
Coonabarabran	2	61,711	6	50	927.2	130	0.140	0.15	0.22
Griffith North	2	65,758	6	49	948.4	139	0.147	0.21	0.59
Griffith South	2	64,126	7	60	1,067.2	249	0.233	0.25	0.62
Lower Darling	1	56,599	7	56	1,071.0	859	0.802	0.08	0.11
Narrabri	2	65,755	7	53	993.1	289	0.291	0.12	0.22
Tibooburra	1	54,849	7	56	1,123.2	457	0.407	0.12	0.25
Total		555,280	56	453	8,634.0	3984	0.486	0.04	0.08

Proposed method

The proposed method uses a model-based approach to estimate the variance of the encounter rate (see section 6.3.2 Buckland *et al*, 2001), which we implement as a generalised linear model. We model the number of encounters in KMZ i and block j, given the transect length l_{ij} in the block, we model the expected number of detections as

$$\mu_{ij} = \mu_i l_{ij} = e^{\beta_i} l_{ij} = e^{\beta_i + \log(l_{ij})}$$

where μ_i is the expected encounter rate in KMZ i. Our model for the number of detections in block j of KMZ i is then:

$$n_{ij} \sim qP(\mu_i l_{ij})$$

$$E(n_{ij}) = \mu_i l_{ij} = e^{\beta_i + \log(l_{ij})}$$

$$var(n_{ij}) = b\mu_i l_{ij}$$

where b is the dispersion parameter of the quasi-Poisson distribution and also the constant ratio of variance to mean across all KMZs, and $\log(l_{ij})$ is an offset term.

After fitting this model to the observed numbers of encounters and line lengths, estimates of the expected values, variances and CVs for the encounter rate in KMZ i can be obtained from

$$\hat{E}(n_i) = e^{(\hat{\beta}_i + \log(L_i))}$$

$$\widehat{var}(n_i) = \hat{b}\hat{E}(n_i)$$

where $n_i$is the number of detections in KMZ i, and $L_i = \sum_i l_{ij}$ is the total line length in the KMZ. The CV of the number groups is given by:

$$\widehat{CV}(n_i) = \frac{\sqrt{\widehat{var}(n_i)}}{\hat{E}(n_i)}$$

Results

The estimate of the dispersion parameter, b, found from the model fitted to all data was 31.02; empirical values are shown in Figure D1. The model-based encounter rates, variances and CVs for each KMZ are shown in Table D2. These results show a mixed picture with some CVs very similar between the two methods, some CVs have decreased and some increased (Figure D2). The decrease in CV with the model-based method occurred in KMZ where there were large, empirical estimates of b (i.e., Griffith North and Griffith South).

When the data were divided into two groups, the estimates of the model-based dispersion parameter were 21.05 and 42.35 in group 1 and 2, respectively. In comparison to the model-based estimates of CV overall groups, the model-based estimates by group have reduced in group 1 and increased in group 2. This is perhaps not surprising but probably better reflects the variability in each KMZ.

Table D2. Model-based estimates using all the data and also the CV estimated by group and, as a reminder, the design-based CV.

KMZ	Group	$\hat{E}(n_i)$	$\widehat{var}(n_i)$	$CV(n_i)$	$CV(n_i)$ Group	Blocks $CV(^n/_L)$
Bourke	1	204	6327.98	0.39	0.32	0.25
Broken Hill	1	1485	46063.96	0.14	0.12	0.13
Cobar	1	172	5335.36	0.42	0.35	0.41
Coonabarabran	2	130	4032.54	0.49	0.57	0.22
Griffith North	2	139	4311.71	0.47	0.55	0.59
Griffith South	2	249	7723.86	0.35	0.41	0.62
Lower Darling	1	859	26645.75	0.19	0.16	0.11
Narrabri	2	289	8964.64	0.33	0.38	0.22
Tibooburra	1	457	14175.91	0.26	0.21	0.25

Figure D1. Distribution of the number of groups per block in each KMZ. The numbers along the bottom are empirical estimates of b.

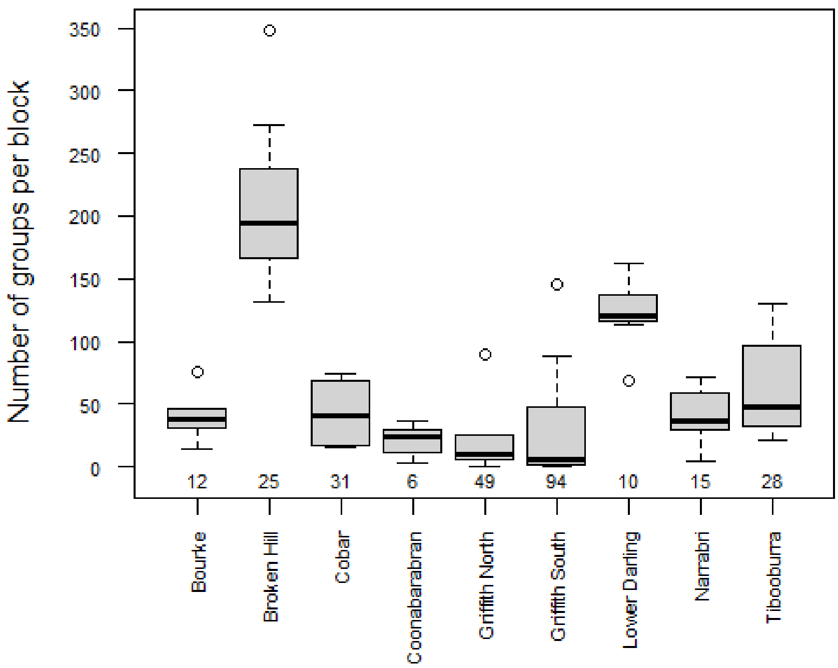


Figure D2. Plot of empirical and model-based CVs for each KMZ: a) overall KMZs and b) by KMZ group. The dashed line is the line of equality.