CLARIFICATION OF ANSWERS GIVEN BY DON FLETCHER AT THE HEARING ON 27 MARCH 2024

At the hearing on 27 March 2024 I was asked several questions about differences between counting methods, and about the potential for comparing methods. The topic is complex and my off-the-cuff responses may have been confusing, so I hope this clarification will be of assistance. The questions and my responses occupy pages 18 to 24 in the transcript of my evidence. Most of the questions were from Ms Abigail Boyd and there were some also from Mr Wes Fang (p 20) and Ms Sue Higgison (p 23). This clarification is six pages long, compared to seven pages for that part of my evidence.

Table 1 is the most important part to read. However the text preceding Table 1 provides necessary explanation. **So it may assist to view this document either on a wide screen or printed single sided, so that Table 1 (page 3) can be examined at the same time as reading the other pages.**

The areas to be counted

Before 2019, most counting of feral horses was applied to the entire Australian Alps National Parks system (**AANP**), of which Kosciuszko National Park (**KNP**) is the largest part. At 16,500 sq km in area the AANP is one quarter the size of Tasmania. Fortunately, the feral horses occur in discrete blocks, which can be counted individually. Even so, very few counting methods can 'scale up' to the extent necessary. Methods that obviously could not scale up sufficiently include counting from horseback or filming with a recreational drone, yet both have been used for 'independent' counts. All 'independent' counts have been of small areas, i.e. under 400 sq km.

Twelve methods for counting horses

Twelve different counting methods have been attempted for feral horses in KNP. And some methods have been referred to by a number of different names. The five methods I am aware of being discussed in early 2024 are summarised in Table 1.

Inherent limitations of methods

The most important distinction between count methods is between:

- indexes of abundance (green column in table); and
- measurements of absolute abundance (yellow column).

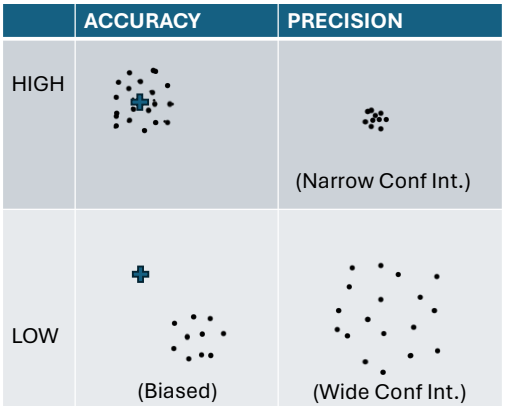
Indexes do not estimate how many individuals were not detected. Indexes do not provide a measure of population size. If an index is repeated using exactly the same procedure, eventually over time it can be used to measure population growth rate and the effectiveness of any population reduction attempts.

Measurements V estimates

In this document, I have deliberately used the term 'measurement' rather than 'estimate'. Because some ecologists refer to the results of their counts as 'estimates', lay people sometimes mistakenly undervalue those counts in favour of other counts that use different terms (exemplified in the hearing on 25/2/24). But all measurements without exception, even micrometer measurements of highly accurate machine parts, fall within a confidence interval. Many measurements also have a degree of bias (inaccuracy).

Terms about measurements: accuracy, bias, precision, confidence interval and mark-recapture

The terms '**bias**' and '**accuracy'** can be understood most easily by analogy to target shooting as shown in the diagram. Both terms refer to the distance from the centre of the target (cross symbol) to the centre of the group of shots (repeats of the measurement). Continuing the analogy, the terms '**precision'** and '**confidence interval**' both refer to the size of the group.



A potential source of confusion is that names of two of the five counting methods include the words '**mark'** and '**recapture'**, yet no horses are marked or captured. Both methods arise from the thinking behind another wildlife survey method in which animals are captured and marked (e.g. by fin clipping of fish). The proportion of recaptures in subsequent capture events can be used to reveal the number of the entire population, including the animals never seen. In the **Helicopter Mark Recapture** method, each horse is photographed and described. When recognised in later surveys these are deemed to be 'recaptures'. The other method is very different. In **Mark Recapture Helicopter Line Transect Distance Sampling** two observers on one side of a helicopter both record groups of horses. The number seen by the rear observer which were also seen by the front observer, are deemed to be recaptures. So then the estimation of the proportion of horses not seen, is based on a combination of mark recapture analysis and distance sampling analysis, which is more accurate.

Importance of stating the confidence interval

With all counting methods, it is fundamental to quantify the precision by estimating a confidence interval around the index or measurement. Without that, there is no way to evaluate reliability, i.e. how different the result might have been if the count was repeated, and so there is no way to know if different counts reflect real difference, or just random variation. For example if you were told the amount of gold in a mine was high enough, you might feel like rushing to invest in it, but if you also learned that the 75% confidence interval for the measurement of the gold was wide enough to include zero, you might not commit. There is a popular edict that 'all biological measurements must specify a confidence interval'.

Estimation of the confidence interval is built in to the three measurement methods mentioned in Table 1 but has not been carried out for either of the index methods i.e. no confidence interval is provided for either the **Horse Helicopter Index** conducted by NPWS staff, nor the **Air Photo Index** provided by Rocky Harvey, Claire Galea and Airborne Logic. These are important gaps.

Its ecology, not mechanical engineering

All measurements of abundant wild populations in large natural areas are somewhat imprecise, even if the author has failed to state a confidence interval. Also, once a certain level of precision is achieved, further improvement usually proves practically impossible even if the counting budget is increased greatly. That is, perfect counts are unattainable.

Table 1: Horse counting methods used in KNP which have been attracting discussion recently.

Indexes of abundance	Measurements of Absolute Abundance
Air Photo Survey (APS) Only one APS has been used in KNP, i.e. the **2024 AirBorne Logic survey** designed by Claire Galea and Rocky Harvey. An airphoto was made of ~161 sq km of open plains and ~51 sq km of adjoining wooded area (212 sq km in all) from high above. More than 405 horses can be seen in the 161 sq km. The APS method lacks any way to estimate a confidence interval. It also lacks any way to estimate the proportion of horses not seen. Compared to HHI, the method detects far fewer horses (Figure 1) because of time of year, time of day, and because the horses are not seen to move. The method is by far the most expensive ever attempted in the AANP, equivalent to about $1.5M for the area which HLTDS covers at less than one tenth the cost.	**Helicopter Line Transect Distance Sampling (HLTDS)** Carried out 9 times from 2001 to 2023 by different groups of workers, of which Cairns (2023) is the latest. (The full list of counts is in Appendix 2 of my submission). The result of the latest measurement was that the KNP feral horse population is 17,432 (95% CI 12,934–22,536). Population growth was about 17% per year between the two big bushfires and 15% including the 2020 fire. (Figure 2) This and MRDS are the only methods ever to have estimated the horse population of either KNP, or the AANP. Both methods also appear to be the only ones with potential to do so..
Horse Helicopter Index (HHI) The HHI is a manual count of all horses seen from a helicopter flying at a low level all around the open plains. Movement by the horses increases the proportion that are detected by the human eye but the method lacks a way to estimate the number of unseen horses. The HHI was carried out annually in September, by select NPWS staff, fourteen times from 1998 to 2021. The latest index counted 3,699 feral horses in the open and the combination of counts shows steady population growth on the plains of almost 14% (Figure 1). The method was not conducted in a way which enabled an estimate of a confidence interval. However it is better and cheaper than APS.	**Mark Recapture Helicopter Line Transect Distance Sampling (MRDS)** Used only once so far in KNP. See Laake *et al*. 2008. Found to be more accurate than HLTDS, which underestimated. Requires a special seating arrangement in helicopters and so perhaps a larger aircraft. MRDS and HLTDS are the only methods attempted in KNP which have capability to survey the entire horse population of KNP. **Helicopter Mark Recapture (HMR)** Requires each horse to be recognised individually when resighted in subsequent surveys, thus limited to small populations. Used only once in the AANP to count 89 (± 10.4) to 94.7 (± 15.5) horses on Bogong High Plains (Dawson and Miller 2008). In the hearing on 18/12/23, Claire Galea proposed using this method, but it has not been attempted. This could be the best method to use in the smallest retention areas after they are culled but is unsuitable for north Kosci.

Which is more important, precision or accuracy?

People who seek a 'head count' of wildlife, in my experience, are always thinking wrong, as if wildlife were like cutlery in a drawer. As stated earlier, the perfect count is unattainable. The appropriate question is about whether the count information is **adequate** to enable management decisions to be made with **sufficient** confidence. To answer that question, it is necessary to recognise when greater accuracy is more important than greater precision, and vice versa. It is also necessary to recognise when it is wiser to use an upper or lower confidence value in preference to the central value (such as the **75%, 90%, 95% or 99% confidence value**). In general:

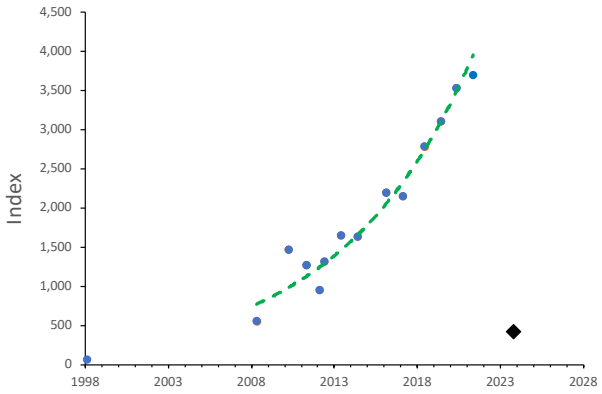
When planning action to save a population from unwanted decline (e.g. a threatened species) consider working from one of the lower confidence values of its population measurement.

When planning action to reduce the population of an invasive species, consider working from one of the upper confidence values of its population measurement, e.g. the upper **90%** confidence value. Note that this is opposite to what Mr Fleming said that NPWS will do in 2024. Presumably his use of the **lower 95% confidence value** is arising from an abundance of caution to avoid over culling. In the next three years it is important for NPWS to gain confidence in all aspects of horse counting and horse management, in order to achieve the required population size in each area.

The Horse Helicopter Index (now discontinued)

The Horse Helicopter Index grew at almost 14% per year with remarkable consistency (Figure 1). The variable efforts put into horse population control during those years had no effect. It show that all those years of trapping, rehoming, and sending horses to the knackery should not have been allowed to continue because the animal welfare cost was imposed on the horses for no benefit. Likewise for the financial cost and the political cost.

Figure 1: Circles are the results of the Horse Helicopter Index (HHI) conducted from 1998 to 2021 on the northern open plains. The dashed green line represents 14% annual growth. The diamond is the first and only Air Photo Index (API) conducted over ~40% of the plains area in February 2024. Note the lack of error bars on all data points.



Population theory predicts that in a population whose distribution is expanding, as the KNP horse population is, population growth rate will be lower than average in central areas which have been occupied longer by the population. This is the case with the 14% population growth rate on the open plains (Figure 1) compared to the 17% average growth of the entire population (Figure 2).

Combining or comparing different count methods

One specific question I was asked repeatedly, was about evaluating the HLTDS (carried out by contractor Stuart Cairns) by comparison with the APS (carried out by contractor AirborneLogic). The latter is an index and indexes can not be compared quantitatively with each other, nor with measurements of absolute abundance such as HLTDS. It is guaranteed that the HLTDS measurement will be higher than the APS one, no matter whether NPWS has culled many horses or no horses. So nothing valid can be learned by comparing between the two. And they cannot be combined. The comparison is invalid primarily because the APS has no way to estimate the proportion of unseen horses. When I last saw it, the APS count was over 405 horses seen and still increasing. But for all anyone knows there could be anwhere from zero to quite a few thousand more horses that have not been (can not be) counted with the method.

In spite of that, this comparison has already been carried out on pro-brumby social media pages and by Rocky Harvey and Claire Galea in a You Tube video. And I know that, whatever I think, the (invalid) comparison will continue to be made, so it could be instructive to consider not just the one popular hypothesis to explain it, but as many hypotheses as seem reasonable. That is, include the explanations (aka hypotheses) proposed by brumby lobbyists, but also consider hypotheses based on the characteristics of the methods and sites, as follows. So, the difference between the results of the methods could hypothetically be because either:

1 As brumby lobbyists claim, the HLTDS has always been exaggerating the population measurement. (It would have to be more than a 10 times exaggeration to explain the difference);
2 As brumby lobbyists claim, NPWS has secretly removed a large number of feral horses from the Northern Kosciuszko Survey Block since last October. (It would have to be around 12,000 horses removed);
3 The APS method lacks any way to estimate the proportion of horses not detected;
4 The APS counted only 41% of the open plains;
5 The open plains are only 30% of the Northern Kosciuszko Survey Block;
6 The 41% of open plains counted by the APS had a lower density of horses than average on the day of the survey, due to disturbance in preceding weeks and during the survey, including by the construction activity which was underway.

A rigorous statistical evaluation of the hypotheses is not possible. My experience tells me that hypothesis 3 would prove to be the most important, that hypotheses 4 and 5 are also true and that hypothesis 6 is likely to be true to some extent.

So what should be done?

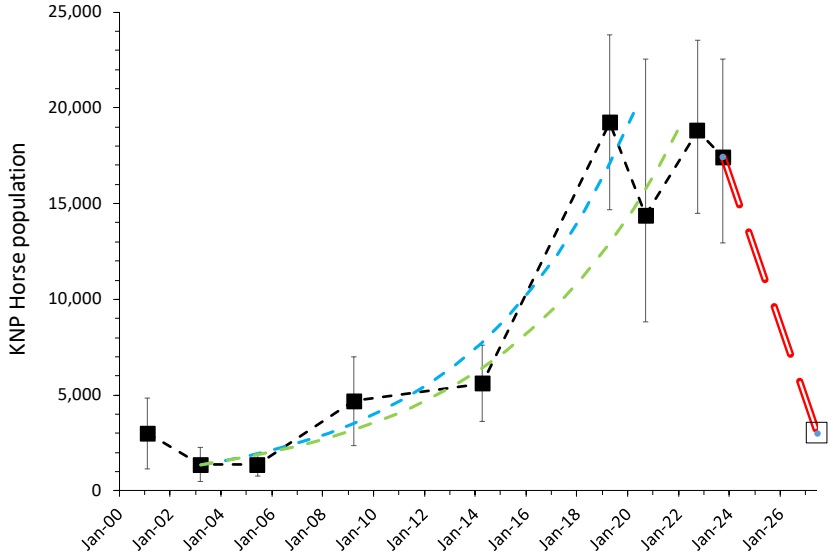
The fundamental requirement is to accept the scientific advice that the best data currently available on horse abundance are those from the HLTDS (Figure 2).

In Figure 2, blue and green dashed lines show what would have happened if the population left after the 2003 bush fire had grown at 17% and 15% per year respectively. The actual counts appear to follow these trends. The 15% rate includes the effect of the 2020 bush fire, so the 17% is probably more indicative of what to expect in the next few years. This is valuable to know, because a culled horse population is likely to grow at this rate, or higher. This means that if a population of 3,000 was culled annually, 450-510 horses have to be removed every year to maintain the population at less than 4,000. Because this inquiry is about **animal welfare**, it is worth noting that if culling was neglected or postponed until the tenth year, more than twice as many horses have to be removed.

This would be unacceptably inhumane in my opinion, because suffering equals the number of animals affected, times the average experience of each animal, so whether an intervention is of a kind that is relatively humane, or one that is relatively cruel, it has double the animal welfare cost if the number of animals is doubled. The red dashed line in Figure 2 represents the change required by law. Also for animal welfare reasons, the faster the population is reduced toward 3,000 the more humane it will be because fewer animals will be impacted. Therefore, for the same reason, it would be cruel to either delay the reduction or to slow it down.

Future counts have to overcome the challenge of estimating lower feral horse densities in each retention area. As mentioned in Table 1, there is likely to be a role for Helicopter Mark Recapture in the smallest of the retention areas. In the larger areas, Mark Recapture Helicopter Line Transect Distance Sampling is advised to replace Helicopter Line Transect Distance Sampling. Thinking outside the square, there could also be good value from using GPS tracking to gain information about horse movements.

Figure 2: Solid squares (± 95% CI) are the results of Helicopter Line Transect Distance Sampling (HLTDS) over all four horse survey blocks in KNP from 2001 to 2023. See text for explanation.



CONCLUSION

- **The best data on abundance of feral horses are the results of the HLTDS (Figure 2).**
- **For the future, consider HMR for the smallest populations and MRDS for the others. Also consider GPS tracking research;**
- **The recent APS demonstrated technical excellence based on ecological ignorance. It failed the most important requirements of a count method.**

REFERENCES

Cairns, SC. (2023). A survey of the wild horse population in Kosciuszko National Park, October 2023. Report to the National Parks and Wildlife Service, NSW. Available from https://www.environment.nsw.gov.au/topics/animals-and-plants/pest-animals-and-weeds/pest-animals/wild-horses/managing-wild-horses/kosciuszko-national-park-wild-horse-management/tracking-the-wild-horse-population

Dawson, MJ. and Miller C. (2008). Aerial mark–recapture estimates of wild horses using natural markings. *Wildlife Research* **35**, 365–370.

Laake, J. Dawson, M.J. and Hone, J. (2008). Visibility Bias in aerial survey: mark-recapture, line-transect or both? *Wildlife Research* 35: 299-309.