SC/66a/SM/15

How many Hectors dolphins off the east coast of New Zealands South Island?

Elisabeth Slooten

INTERNATIONAL
WHALING COMMISSION

How many Hector's dolphins off the east coast of New Zealand's South Island?

Elisabeth Slooten, University of Otago, P.O. Box 56, Dunedin, New Zealand

ABSTRACT

A recent survey of Hector's dolphins off the east coast of New Zealand's South Island was reviewed by local and international experts. Several serious problems with survey design and data analysis were identified, including problems with the estimation of availability and perception bias, in particular the analysis of circle-back data which did not consider dolphin movement. Covariates such as glare and weather conditions were not included in the analysis. An unusually large amount of data were removed from the dataset before analysis by truncation: 34% of the data from all four observers and 42% of the data from the two observers who were able to see the trackline. The assumptions of the analysis method (Mark Recapture Distance Sampling) were violated. Several models that appeared to fit the data reasonably well produced abundance estimates much higher than the other models, which were described as 'unrealistic' by the researchers who carried out the survey. In one case the estimate of the number of dolphins within the survey strip was 1.5 million Hector's dolphins, indicating a total abundance estimate of around 35 million Hector's dolphins off the east coast of the South Island. Instead of investigating why some of these apparently well-fitting models produced abundance estimates of several million dolphins, the estimates from these models were simply left out of the model-averaged abundance estimates, while including estimates from models that apparently fitted the data less well. The implications of the survey results for the potential impact of bycatch on this population are explored using an agent-based population model. The results indicate that the rate of population increase/decline is relatively insenstive to population size but highly sensitive to estimates of fishing effort and catch rate. This has important research and management implications.

INTRODUCTION

A population survey of Hector's dolphins off the east coast of the South Island of New Zealand (NZ) was carried out during January-March 2013 (austral summer) and July-August 2013 (winter), under contract to the NZ government (MacKenzie and Clement 2014a). Government invited peer review of the survey results from Prof Phil Hammond and Prof Steve Buckland (University of St Andrews) and two meetings with stakeholders (in May and October 2013) including representatives of the fishing industry, conservation groups and government as well as a small number of local population survey experts (Prof Steve Dawson, Prof Liz Slooten and Dr Will Rayment from University of Otago were able to attend some of these meetings). In addition, Dr Karin Forney (US National Marine Fisheries Service) provided a separate review, which was circulated as an informal working paper at the 2014 Scientific Committee meeting. The appendices to MacKenzie and Clement (2014a) can be found in a second document, which includes the authors' responses to the peer review (MacKenzie and Clement 2014b).

This paper briefly summarises key problems with the survey design and data analysis, identified by reviewers but not yet corrected by MacKenzie and Clement (2014a,b). In addition, the implications of the survey results for the potential impact of bycatch on this population are explored using an agent-based population model.

PROBLEMS IDENTIFIED BY REVIEWERS

Detection functions

The detection functions for the summer and winter surveys are shown below in Figures 10 and 14 from MacKenzie and Clement (2014a). The summer detection function peaks at around 100-200 m from the trackline. In contrast, the winter detection function is highest on the trackline and is relatively flat from the trackline to a distance of about 120 m. Examination of the sightings data themselves shows that observer behaviour changed dramatically between the summer and winter surveys. The substantial difference between summer and winter in the distribution of sightings from the rear observer indicate that the observers spent much more effort ensuring a high sighting rate on the trackline in winter and therefore less effort in the overlap zone between front and rear observers. As a consequence, the the proportion of duplicates was reduced, and the overall result was a lower abundance estimate (compared to summer).

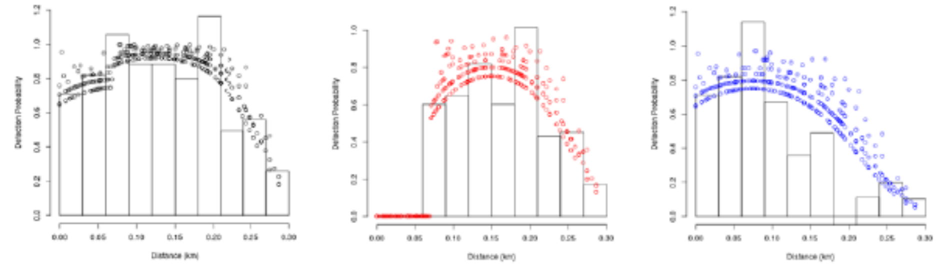

Figure 10: Fitted detection functions and histograms of empirical detection probabilities from the top ranked model in Table 5. Left is $p_{\bullet}(d_i, s_i)$, centre is $p_F(d_i, s_i)$, and right is $p_{R|NF}(d_i, s_i)$.

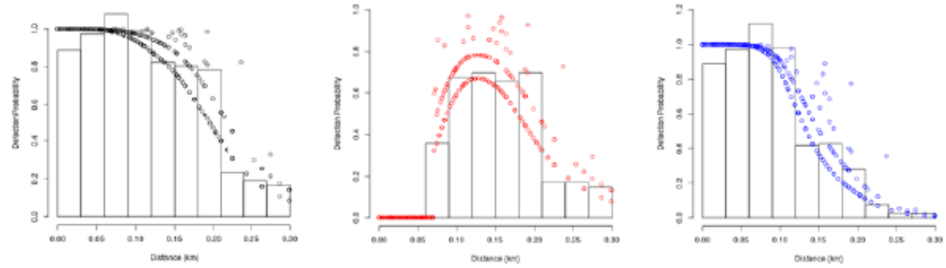

Figure 14: Fitted detection functions and histograms of empirical detection probabilities from the top ranked model in Table 11. Left is $p_{\bullet}(d_i, s_i)$, centre is $p_F(d_i, s_i)$, and right is $p_{R|NF}(d_i, s_i)$.

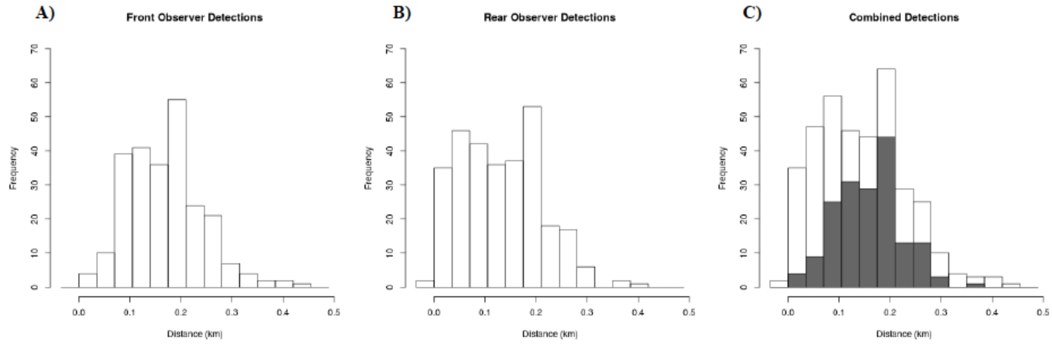

Figure 4: Histogram of the verified distance data prior to any truncation of sightings from A) the front observer position, B) the rear observer position, and C) from either observer position in the summer survey. Grey bars indicate sightings made by both observers (duplicates).

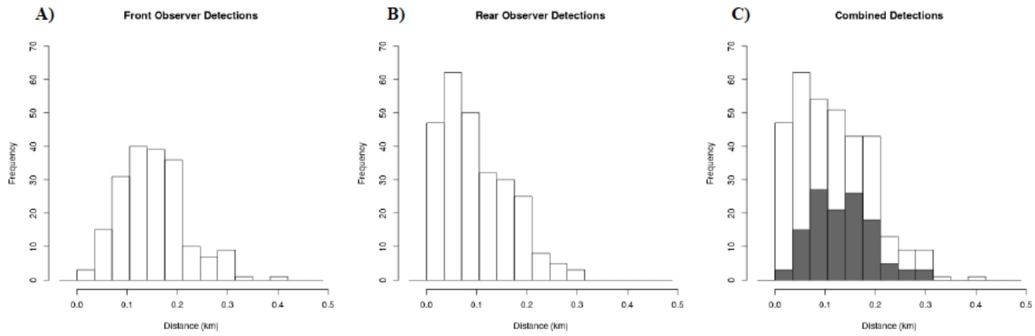

Figure 7: Histogram of the verified distance data prior to any truncation of sightings from A) the front observer position, B) the rear observer position, and C) from either observer position in the winter survey. Grey bars indicate sightings made by both observers (duplicates).

Further analysis may help to quantify the effect of the change in observer behaviour on the abundance estimates and determine if the difference in the summer and winter abundance estimates reflects methodological differences, a real reduction in dolphin densities in winter or a combination of both. Another major difference in the field methods between summer and winter concerns the line spacing. Effort was removed from Cloudy-Clifford Bay and added to Banks Peninsula. Again, the effect of this change in survey effort on the abundance estimates for these two areas, and the overall ECSI estimates, is not clear. Further analysis would be required to quantify these effects. The survey data have not yet been provided to government. Once this is done, government has undertaken to make the survey data publicly available, which will make it possible to quantitatively explore these issues.

Other problems with survey design

Several of the survey lines run alongshore, which is contary to the usual sampling practice of placing lines across known gradients of density. Density declines with distance offshore, and varies alongshore. Therefore previous surveys have used transects are 45^{o} to the shore. The maps of sighting locations indicate that this resulted in higher sighting rates in summer. In winter, dolphins were less concentrated close to shore – reducing the impact of this survey design problem. Prof Hammond commented that the program Distance should have been used to design the survey which would have ensured equal coverage for inshore and offshore areas.

Availability and perception bias

A high level of uncertainty about perception bias and availability bias is reflected in the widely ranging estimates of Effective Strip Width (ESW) and highly variable abundance estimates produced by different models. For example, estimates of availability bias ranged from 0.25 to 0.93.

Model-averaged estimates for large areas (e.g. north and east coast of the South Island combined) understate the level of uncertainty associated with the estimates. Both international peer-reviewers commented on the ad-hoc manner in which duplicate sightings were identified. This was a problem in the circle back experiments as well as the normal survey sightings, and was also discussed at the two local stakeholder meetings.

For example, Appendix Q of MacKenzie and Clement (2014b) includes the following exchange:
‘Prof. Hammond’s comments and our responses
Main points:
PH:
The designation of duplicate sightings in the main survey is not well justified. Why 5 seconds, 5 degrees and what was the group size tolerance? How sensitive were the results to variation in these threshold values? This could make a considerable difference to the perception correction.

DM/DC:
a range of criteria were used to decide whether a sighting was a duplicate sighting of a group already recorded by the other observer, including time between sightings, recorded angles, group size and observer comments. 5 seconds and 5 degrees were used as guidelines, but were not strictly adhered to with experience playing a leading role in the process (note that all matching was done manually).’

MacKenzie and Clement (2014b) agree that ‘Clearly there is the potential for bias to be introduced through the misidentification of duplicates, and in some circumstances the bias can be quite extreme.’ However, they then ignore the problem, arguing that ‘biases may be either negative or positive, and given the protocols used during the surveys, in duplicate matching and DC's [Deanna Clement’s] experience in aerial survey work for Hector's dolphin (dating from the early 2000s as an observer on surveys on the South Island's West Coast), it is hard to imagine misidentification rates in excess of 5%, especially given the low frequency of any dolphin sightings during the surveys.’

MacKenzie and Clement (2014b) comment that 'The process of duplicate matching was done manually and as such is not a simple task'. Of course this does not ensure that the process was robust or consistent. As with most of the other peer-review comments, MacKenzie and Clement (2014b) failed to respond, arguing either that to investigate the problem pointed out by the reviewer would take a substantial amount of extra work and/or that biases were likely to be small or likely to cancel each other out. For example on the topic of duplicate identification, MacKenzie and Clement (2014b) responded that 'Investigating what effect slightly different criteria may have on estimated abundance is therefore a substantial undertaking. Given that … any effect is likely to be relatively minor and not practically alter the overall conclusions (particularly when the width of the confidence intervals are considered), we do not believe a reworking of the data and reanalysis is worthwhile at this stage.' Typically, the only argument for failing to correct or examine problems identified by reviewers is a lack of time or resources.

A repeated comment in the peer reviews was that the analysis appears to have over-corrected for perception bias. The MRDS analysis corrects for perception bias; the circle back method attempts to estimate availability and perception bias; the helicopter studies address availability bias. Thus combining MRDS with the circle-back method would be double-correcting for perception bias, unless only one team's data were used (which does not appear to be the case). The data needs to be re-analysed in order to determine how these different estimates were combined and to ensure that there is no over-correction.

Further, if the circle-back method of estimating Pα (which includes both availability and perception bias) is used, then Eq (10) on page 21 (MacKenzie and Clement 2014a) will result in an overestimation of abundance. That equation is correct if the availability correction factor is developed for animals on the transect line (perpendicular distance $y = 0$). However, the circle-back method in this study appears to have estimated Pα for the entire viewing area (of both observers), which means it also includes distance-related effects on detection probability at some distance y. However, the distance-related effect on detection has already been accounted for in Equation (8) on page 20. Therefore, applying the circle-back correction in Eq (10) double-corrects for the distance-related effects. This may be further compounded by the fact that the peak detection probability is not at the (left-truncated) zero distance and differs by season, but insufficient details have been provided to evaluate the magnitude of these issues.

Several reviewers commented on regional differences in availability bias. For example, Prof Hammond's comments below:
'PH: P44: Availability estimates from dive cycles
This regional variability seems quite high given the SEs. Is there any information about what could cause this? Presumably methods used to collect the data were the same in all regions.'

MacKenzie and Clement (2014b) speculate that this may be due to regional differences in environmental factors, but do not explore the potential impact on the abundance estimates due to lack of time and resources:
'DM/DC: consistent methods were used in all regions. The effects may be due to differences in water depth, temperature or local ecosystems. Understanding why such differences might exist, however, was outside of the contract and have not been explored.'

Equation (6) on page 17 does not include sample size (# groups) in the estimation of the standard error. Sample size should be part of this calculation. The Kaikoura winter estimate, based on a very small sample size of only 6 groups (Table 16 in MacKenzie and Clement 2014a), is very different from the other estimates that were based on larger sample sizes. This suggests imprecision due to small sample size, and probably over-stratification when estimating availability.

The similarity in durations of the circle-back procedure (5-10 sec + 20 sec + 60-120 sec = 85-150 secs) and the dive cycle of Hector's dolphin (~70-137 sec, Table 15) is a concern. This means the detection of animals on the second pass is not independent of the first detection, an important assumption of this method. The choice of the circle-back timing should be examined, to determine whether it was adequate and unbiased. MacKenzie and Clement's (2014a) statement 'However, given the degree of natural variation in the time taken to complete circle-backs…and in Hector's dolphin dive cycles, the point in the dolphin's dive cycle at which the aircraft files over the group could be considered random' is unconvincing without supporting evidence. Laake and Borchers (2004) and others have discussed the importance of choosing an appropriate time interval. This is all the more important given that it is usual practice for pilots to judge their turns using the 2-minute turn mark on the 'turn co-ordinator' instrument. This is known as a standard rate turn, and completes a 360^o turn in two minutes.

Movement of animals out of the viewing area during circle-back protocols needs to be considered in the analysis. The total viewing area is 229 m wide (from 71-300 m). Dolphins traveling at a moderate speed of 4 knots would travel about 250 m in the time it takes to circle back, making it likely that a non-trivial proportion of animals were missed because they were no longer in the strip, rather than because of perception/availability bias. In any case, the animals would almost certainly have moved more than the 5 degree tolerance for re-sightings (recognizing that this was only one factor used). MacKenzie and Clement (2014a) state they had no way of estimating the rate of movement. This was clearly a major oversight in survey design. As soon as the choice was made to use circle-back corrections, estimating he rate of dolphin movement should have been included in the fieldwork component of the project. In fact, there is ample existing information on movement rates from boat-based research. Even without using those data, it would be relatively easy to assess whether the dolphin groups included in the circle-backs were traveling slowly, rapidly, or milling during surveys. If no such information was collected, then uncertainty around circle-back corrections must include this major source of uncertainty – perhaps using simulations in which varying proportions (e.g. 20%, 50%, 75%, 90%) of animals move away and are no longer in the survey strip.

Unmodelled heterogeneity can cause serious problems in circle-back analyses (e.g. Thomsen et al. 2005), and may have played a role in the very high abundance estimates produced by some of the models. For example, the estimate of 1.5 million dolphins in Table 7 (for more information, see section on 'unrealistic' abundance estimates below).

The choice of t = 6 seconds (for the amount of time an object is within the field of view of an observer, see MacKenzie and Clement 2014a pages 43-45) seems a bit long for surveys flown at 500 ft and 100 kts, even at a 65 degree declination angle. In aerial surveys at 600-700 ft and 90-100 kts (which gives a longer effective viewing time), animals can effectively be detected at 65 degrees for only ~2-5 seconds (Forney pers. comm.) depending on cloud cover, sea state, glare, and whether the animal is surfacing or submerged.

In the following exchange, MacKenzie and Clement (2014b) agree with Prof Hammond's critique of their circle-back approach, but do not act on his suggestion to include dolphin movement in the analysis.

'PH: In the circle-back analysis, it is assumed that groups stay in the survey strip; however, some may move out of view in the time taken to circle back (a few minutes). This could potentially cause resightings to be missed, availability to be underestimated and abundance to be overestimated. Uncertainty in resighting is referred to but otherwise apparently ignored in analysis. It is not clear that availability as determined from the helicopter dive time experiments is the same as availability on the surveys themselves. This is critical to avoid bias in the use of this correction. Bias that could potentially result could be in either direction.

DM/DC: we agree wholeheartedly with your comments regarding availability measures and their shortcomings. We are fully aware of the issues you raised, have voiced similar concerns ourselves during meetings and long noted that there is insufficient emphasis on making resources available for studies into a robust method for estimating availability. Indeed, it was concerns about how accurately availability estimated from helicopter surveys of dive cycles hovering near groups (as had been previously used) reflected availability for observers in fixed-wing aircraft flying transects that led us to trialling, and implementing, the circle-back methodology. Given the resources available these were the best, but imperfect, options open to us. That both methods give similar results is encouraging, although it's impossible to determine whether both are ok, or both are badly wrong, without information from a better method for assessing availability. While you have noted there is the potential for biases, you do not appear to be making any clear recommendations on what could be done differently to improve the analysis given the data that is available.'

The simple solution to the problems pointed out by the reviewers, is to use standard analysis methods that take into account animal movement. The 'clear recommendations' for a 'better method' are already available in the literature on line-transect surveys. The circle-back data should be re-analysed using standard analysis protocols (e.g. Hiby 1999; Thomsen et al. 2005) replacing the ad-hoc decisions used by MacKenzie and Clement (2014a).

Covariates

Although sea state and glare information was collected during the survey, these important covariates are not included in the analysis. In essence, this assumes there is no effect of these factors on detection and/or no heterogeneity in these factors among strata. It seems unlikely that sea state/wind conditions are the same inshore/offshore and in all areas along the coast, yet the analysis assumes a common detection function across all strata. Prof Buckland commented that low sample sizes in areas of low density can force a common detection function, and that this is less important when covariates are included in the analysis. MacKenzie and Clement (2014b) responded with the statement 'a common detection function was assumed regardless across all strata given consistent field methods were used'. The issue is not whether consistent field methods were used, but whether environmental factors (e.g. water clarity, sea state) and biological factors (e.g. dive depth, time spent at the surface) were consistent across all strata.

For small cetaceans, sea states in the range Beaufort 0 to 3 are well-documented to markedly affect detection rates, especially for small species like Hector's dolphins. Further, from personal experience conducting extensive aerial surveys it is clear that cloud cover (above the plane) also dramatically affects the ability to see animals below the surface. In overcast conditions, even an obvious humpback whale or strikingly marked *Lagenorhynchus* can completely disappear as soon as they sink below the immediate surface. By contrast, in sunny conditions animals can be seen at least a few meters below the surface. Yet this important factor was not mentioned (e.g. were surveys only flown in sunny conditions?). There is no indication that these factors were accounted for in the analysis or calibration.

Information on the proportion of effort by sea state, cloud cover, etc (preferably by stratum) during the surveys and the calibration studies is required to evaluate this potential problem. If these proportions are not constant, then the analysis needs to account for this heterogeneity; otherwise the detection function and calibration factors for availability and perception bias could be highly inaccurate. For example, if the stratum-specific calibration factors were not estimated for similar conditions to those during the surveys within that stratum, they will be biased – potentially by large amounts – in the individual strata.

Truncation

All sightings data were right-truncated to a distance of 0.3 km. After right-truncation, the dataset is described as the 'full' dataset by MacKenzie and Clement (2014a). This is a relatively strong right-truncation, but not out of the ordinary compared to other line-transect surveys for dolphins.

The additional left truncation, at 0.71 km, applied to the 'reduced' dataset indicates a serious problem with survey design. Left and right truncation, leaving only those sightings between 71 and 300 m from the plane, resulted in removing about a third of the data (34% for both observers, and 42% of the data from the rear observer) from the analysis. If both of the observers had been able to see the trackline, it would have been unnecessary to eliminate such a large number of sightings from the analysis. This is particularly problematic given that the shape of the detection functions indicated problems with the field protocol followed by observers.

Violations of the assumptions of Mark Recapture Distance Sampling (MRDS)

The assumption that both observers can see the trackline was violated. MacKenzie and Clement (2014a) developed an extension to MRDS 'due to the survey design where not all distances are observable from the two observer positions' (MacKenzie and Clement 2014a). Appendix B in MacKenzie and Clement (2014b) explains that 'Theoretically this is not a huge extension, although common methods of analysis (e.g., Program DISTANCE and the MRDS *R* package) are not designed for such a situation hence custom code needed to be developed'.

A second extension to MRDS aimed to allow 'potential lack of independence between observers' to be 'incorporated into the modelling'. As explained in Appendix B (MacKenzie and Clement 2014b) 'lack of independence can be caused by observers responding to cues from other observers in the aircraft (e.g., if the rear observer notices movement from the forward observer when they have detected a group, the rear observer may search harder), or due to detection heterogeneity that is unexplained in the model.'

Potential causes for detection heterogeneity in this survey included glare. The survey design meant that many of the lines were flown with glare side-on to the trackline. Both observers on the glare side of the plane would have had low sighting probability, and the observers on the other side would have had high sighting probability on the same line. Other sources of heterogeneity include the fact that some of the observers have much higher sighting rates than others.

MacKenzie and Clement (2014a,b) state that the different approach used in this study to incorporate potential dependence between observers was considered more numerically stable. This is clearly not the case, as some of the models produced very high abundance estimates (see below) and were markedly not stable. Also, it is not valid to assume that dependence among observers exhibits a linear effect when the detection function was unimodal, rather than linear with respect to distance. These analytical assumptions should be thoroughly tested once the data have been made available.

Models that appear to fit the data but produce 'unrealistic' abundance estimates

Some of the models appeared to fit the data reasonably well but produced extremely high abundance estimates, described as 'unrealistic' by MacKenzie and Clement (2014a). For example, in Table 7 (MacKenzie and Clement 2014a) the model with the fourth lowest ΔAIC resulted in an estimate of 1.5 million dolphins in the area 'covered' by the survey (i.e. 2ESW * survey line length), even before correction to estimate the total population size in the area surveyed. In this case, 'model averaging has been performed based upon the models ranked 1-3 and 5-8'. Table 13, likewise includes some very high abundance estimates. In this case, 'model averaging was performed using the top three models'. In general, MacKenzie and Clement's models with relatively high ΔAICs tended to produce relatively high abundance estimates. Therefore, decisions on which models to include in model averaging could have a strong influence on the overall abundance estimate. The importance of these arbitrary decisions on the resulting population estimates is unclear, and will remain so until the survey data are made publicly available and can be re-analysed.

Overall, MacKenzie and Clement (2014a) used a somewhat 'shot-gun' approach, developing many alternate models and then averaging all but the most egregious outliers. While model-averaging can improve analyses, one should first allow the data to guide the analysis as to which models are supported. Failing to do this has added tremendous complexity to the reports (MacKenzie and Clement 2014a,b) without enhancing the quality of the analysis – indeed, making it more difficult to assess the validity of what was done. Fundamentally, the wide range of abundance estimates shown in these models shows that the data were not particularly informative.

Standard distance sampling analysis

At the October 2013 meeting, MacKenzie and Clement were asked to carry out a standard distance sampling analysis on the survey data, to help determine the causes of the problems raised by peer review. A brief appendix in the document that describes the authors' responses to the peer review (MacKenzie and Clement 2014b, pages 74-75) provided only 'a simple comparison to illustrate that the abundance estimates provided by the current methods are realistic'. Without additional information, including the detection functions from the standard Distance analysis, it is still impossible to evaluate problems such as the changes in field protocols, the heavy truncation of the dataset (removing 34% of the data), etc.

As a consequence it is still not clear what caused the high level of variability in the estimates for different models and different seasons. For example, what was the contribution of the unusual shape of the summer detection function and the very different winter detection function?

DISCUSSION OF THE PROBLEMS WITH SURVEY DESIGN AND ANALYSIS:

Many of the concerns summarised above cannot properly be evaluated because the reports do not include the relevant information (e.g. effect of sea state, cloud cover, on results and availability calibrations). In addition, despite repeated requests, the data have not yet been made available.

In the introduction to their response to the peer review, MacKenzie and Clement (2014b) make the following statement about the problems pointed out by the reviewers: 'Overall, while some comments made by both reviewers are valid and could lead to small changes in the estimated abundance for ECSI Hector's dolphins, the resulting differences could be in opposite directions given the aspects the reviewers have noted. Our opinion is that such changes are not warranted at this point given the level of work that would be required to implement them and that they are extremely unlikely to substantially alter the main findings of this project, especially considering the wide range of the confidence intervals.'

The typical response to problems pointed out by reviewers is that it would take a substantial amount of extra work and time, or that biases are likely to be small or likely to cancel each other out. However, no evidence is provided to demonstrate that the biases would be small or would cancel each other out.

For example, the following exchange between Prof Hammond and MacKenzie and Clement (2014b):
'PH: P55: Abundance estimates
There generally seems to be a bigger difference between full and reduced datasets in winter
than in summer. Why should this be? Because there is more variability in the winter
estimates (bigger CVs)?

DM/DC: this aspect has not been fully explored (beyond error checking) as seeking an explanation will not change the results.'

Several of the reviewers commented on the fact that where more than one estimate of group size was available, the higher estimate was used in the analysis. For example, Prof Hammond: 'It is conventional to use mean group size estimates if more than one independent value is available. The decision to use the maximum is not well justified and causes estimates to be higher than they otherwise would be.' Prof Buckland also queried the use of the larger of the two estimates, for duplicate sightings where the two observers differed in their estimate of group size, and commented that 'perhaps that a correction should be used for detections seen by just one observer'. MacKenzie and Clement's (2014b) response was: 'To incorporate any correction properly would require a complete reworking of the entire analysis and lead to a possible change that will be small relative to the width of the current confidence intervals and not substantially alter the overall results. We do not believe it is worthwhile pursuing such a correction at this point in time.'

In fact, there is no reason to believe that implementing the recommendations of the reviewers would lead to 'small' changes in the abundance estimates. Many of the problems with survey protocol, data management (e.g. decisions on which sightings are duplicates) and data analysis concern correction factors (e.g. for availability and perception bias) which could have a substantial impact on the abundance estimates. Bias could be non-trivial for any one of these factors, and considerable if multiple factors are acting simultaneously. MacKenzie and Clement's (2014a) assertion that biases could be in opposite directions and thus '..they are extremely unlikely to substantially alter the main findings..' is unconvincing. The magnitude of each source of bias can vary markedly, and there is no rationale for believing they will all cancel each other out perfectly. Rather, the resulting abundance estimates are far more likely to be biased (upward or downward) than unbiased.

Therefore, it is critical to either 1) correct the problems in the analysis, or, if that cannot be done because data are lacking, 2) estimate the associated uncertainty around these sources of bias and include them in the overall variance estimates. Once these sources of uncertainty are included, the confidence intervals around the abundance estimates are likely to widen considerably.

MacKenzie and Clement (2014a) conclude that 'There is general agreement between our seasonal abundance estimates, confirming that the current population of Hector's dolphin along the ECSI is larger than expected from previous estimates.' (first sentence of Discussion). There are two problems with this conclusion: 1. There are substantial differences between the summer and winter estimates, and 2. Even if the summer and winter estimates were identical, this does not constitute evidence that the population is larger than expected from previous estimates. Both of the 2013 estimates could be markedly biased (e.g. due to the problems outlined above).

The summer estimate was 9,130 (95% CI 6,342-13,144) and the winter estimate 7,456 (95% CI 5,225-10,641). The winter estimate only just fits within the confidence interval of the summer estimate). The summer and winter estimates were sufficiently different for MacKenzie and Clement (2014a) to suggest that this may be due to 'regional alongshore movements and a further than anticipated offshore shift'. Any substantial alongshore movements would be inconsistent with 30 years of intensive photo-ID effort (e.g. Rayment et al. 2009).

The Banks Peninsula area is one of two high density areas on the east coast of the South Island, accounting for more than 5,000 of the individuals included in the summer estimate of just over 9,000. The seasonal difference in the Banks Peninsula abundance estimate is substantial (Table 1, below). The model-averaged winter estimate is 1,999 (95% CI 970-4,330), less than half of the summer estimate of 5,025 (95% CI 2,534-10,552).

	Right truncated		**Left and right truncated**		**Model averaged**		
	Dive cycle	Circle back	Dive cycle	Circle back	N	lower CI	upper CI
Summer	5,458	4,242	5,742	4,656	5,025	2,534	10,552
Winter	2,117	1,939	2,263	1,675	1,999	970	4,330

Table 1. Population estimates for the Banks Peninsula area (east coast South Island) based on right-truncated and left- and right-truncated data, from MacKenzie and Clement's (2014a) Tables 25 (summer) and 27 (winter). The model averaged estimates from Tables 26 and 28, were averaged across the two different levels of truncation and two different methods for estimating availability bias (dive cycle and circle back).

An important consideration, in comparing the 2013 summer and winter estimates or comparing the 2013 estimates with published abundance estimates, is that MacKenzie and Clement (2014a,b) without doubt underestimated the uncertainty in their estimates. For example, Prof Hammond points out that the MRDS analysis is likely to have underestimated variability by failing to include duplicate uncertainty.

On a positive note, some basic conclusions can already be drawn even before the data are made available and re-analysed. For example, the survey indicates that the Hector's dolphin population off the east coast of New Zealand's South Island numbers in the thousands, rather than tens of thousands. It also confirms that most of the population is found in waters less than 100 m deep, with only a few sightings in > 100 m water depth (also see Rayment et al. 2010). Both of these conclusions have important management implications. The offshore distribution information shows there is still substantial overlap between Hector's dolphins and both trawl and gillnet fisheries. For example, off Banks Peninsula gillnets are banned to 4 nautical miles from shore, but Hector's dolphins range to at least 20 nautical miles offshore.

The survey was a compromise between a design suitable for estimating abundance and a design suitable for quantifying alongshore distribution. As a consequence, the results have shortcomings in both regards. The survey confirmed the two high density areas around Banks Peninsula and in Cloudy Bay – Clifford Bay. However, the results for low density areas are inconclusive. For example, in Golden Bay on the north coast of the South Island no sightings were made in summer and one sighting was made in winter. This resulted in an estimate for this area of 0 for summer and 187 (95% CI 32–1,087) in winter. Essentially, this indicates the presence of Hector's dolphins in Golden Bay, which is also known from boat surveys (e.g. Dawson and Slooten 1988; Slooten et al. 2002). It's not possible to conclude anything beyond a relatively low population density in the area, as the abundance estimate for this area is essentially a 'hit or miss' process based on a small number of sightings (e.g. 0, 1, 2 or 3). In another low density area off the Otago coastline, no sightings were made in summer or winter, despite the fact that this area has a known local population of about 37 Hector's dolphins (95% CI 25-75; Turek et al. 2013).

Government has undertaken to make the survey data publicly available (Currey pers. comm. 2013, 2014, 2015). This will make it possible to quantitatively explore the effect of some of the problems outlined above. However, some of the issues appear to be due to problems with the field methods (e.g. the substantial difference in observer behaviour between the summer and winter survey). Therefore, it's possible that even after re-analysis a valid abundance estimate remains elusive.

HOW IMPORTANT ARE ABUNDANCE ESTMATES IN ESTIMATING IMPACTS ON THIS POPULATION?

The importance of population size was explored using a spatially structured, individual based model, and compared with the importance of the number of gillnets and trawling vessels and the catch rates in these two fisheries. The analysis was carried out using Netlogo (version 5.1.0) an agent based modelling platform dowloadable from the Netlogo webpage (Wilenski 1999). The model was parameterized using empirical data from a long-term research programme on New Zealand dolphin (Dawson et al. 2004; Gormley et al. 2012; Slooten and Lad 1992; Slooten et al. 2000) and government data on fisheries mortality (DOC & Mfish 2007; MPI & DOC 2012; Baird et al. 2015; Currey et al. 2012; Davies et al. 2008; Penny et al. 2007). The biological dataset includes population surveys and estimates of movement, survival and reproductive rates. The fisheries data include estimates of fishing effort, catch rates and movements of fishing vessels.

A model was built for the Hector's dolphin population around Banks Peninsula. Dolphins ranged throughout the area, with a preference for inshore waters based on the relative inshore/offshore densities estimated by MacKenzie and Clement's (2014a). Gillnets were stationary and allocated using estimates of fishing effort from the Ministry for Primary Industries. The fisheries data indicate that at any one time there are an average of five gillnets and two trawlers being used in the Banks Peninsula area. Gillnets were placed in a new location each day, to represent typical fishing behaviour. Gillnets are usually cleared every 24-48 hours and then put back in the water. Gillnet and trawling effort were restricted to areas where these fishing methods are legal, i.e. the model does not take account of illegal fishing. Gillnets were placed only > 4 nautical miles (nmi) offshore and trawl vessels only > 2 nautical miles offshore. The modeled dolphin habitat extended to 20 nmi offshore. Trawling vessels were modelled as moving agents, with a movement rate corresponding to the average distance over which inshore trawling vessels deploy their trawl gear (approx. 8 nmi per day; Baird et al. 2015).

In this spatially explicit analysis, the catch rate describes the probability of a dolphin death for any given dolphin encounter with a gillnet or trawl net. While there are no direct estimates of this parameter (e.g. from observers or video camera systems), a bycatch rate of 0.1 for gillnets and 0.01 for trawling is consistent with data from past observer programmes (DOC & Mfish 2007; MPI & DOC 2012; Baird & Bradford 2000; Davies et al. 2008; Slooten and Davies 2011).

Females start to breed at age seven and produced one calf every two years (Slooten and Lad 1991; Slooten et al. 2000). Natural mortality was based on a mortality curve for humans from Slooten and Lad (1992), scaled to a lifespan of 30 years following Barlow and Boveng's (1991) approach. A range of input values for the number of gillnets and trawling vessels was explored, as well as different catch rates for these two fisheries. These values span a credible range of fishing effort, consistent with fisheries information collected by the Ministry for Primary Industries.

	Gillnets															
	5								10							
	Catch rate															
	0.1				0.2				0.1				0.2			
	Trawls															
	2		4		2		4		2		4		2		4	
	Catch rate															
Dolphins	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02
2000	0.99	0.97	0.97	0.95	0.97	0.97	0.97	0.95	0.97	0.97	0.96	0.95	0.96	0.94	0.94	0.93
3000	0.98	0.98	0.98	0.96	0.97	0.97	0.97	0.95	0.98	0.96	0.97	0.96	0.95	0.95	0.95	0.94
4000	0.98	0.98	0.98	0.96	0.97	0.97	0.97	0.94	0.97	0.97	0.97	0.95	0.96	0.95	0.95	0.93

Table 2. Variation in the rate of population decline for the Banks Peninsula area (east coast South Island) depending on population size, the number of gillnets and trawling vessels and the catch rate for these two fisheries.

The goal here was not to provide updated estimates of the total number of dolphins caught or the population's current rate of decline. An estimate of bycatch is available from an observer programme in this area in 1997-1998 (Baird and Bradford 2000) and the rate of population decline has been estimated by Gormley et al. (2012). The goal of this analysis was to explore the sensitivity of the estimated rate of population decline to estimates of population size on one hand and estimates of fishing effort on the other. The results indicate that the rate of decline is not particularly sensitive to to population size but highly sensitive to fishing effort (Table 2).

This has important implications for research priorities. There is currently very little research effort on estimating fishing effort and catch rate. Annual analyses of bycatch for New Zealand marine mammals typically state that it was not possible to estimate Hector's dolphin bycatch due to low observer coverage. For example, this statement from Thompson et al. (2013): 'Hector's dolphin have also been observed caught in set nets in the past, and have also previously been reported as bycatch in trawl fisheries (Baird & Bradford 2000). The incidental captures of endemic Hector's dolphin are of concern, as this species is endangered, with a small population size that is considered to be decreasing (Currey et al. 2012). Hector's dolphin have a coastal distribution, which makes them vulnerable to inshore fisheries, including trawling given the high fishing effort involved. At the same time, low observer coverage in these fisheries precludes reliable estimates of total Hector's dolphin captures.'

These results also have important conservation implications. Essentially, with the same catch rate and fishing effort, a doubling of population size will increase the number of dolphins caught and has very little effect on the rate of population increase or decrease.

REFERENCES

Baird S, Bradford E 2000. Estimation of Hector's dolphin bycatch from inshore fisheries, 1997/98 fishing year. Report for Department of Conservation, Wellington, New Zealand.

Baird SJ, Hewitt J, Wood BA 2015. Benthic habitat classes and trawl fishing disturbance in New Zealand waters shallower than 250 m. New Zealand Aquatic Environment and Biodiversity Report 144, Available from the Ministry for Primary Industries, Wellington, New Zealand.

Currey RJC, Boren LJ, Sharp BR, Peterson D 2012. A risk assessment of threats to Maui's dolphins. Ministry for Primary Industries and Department of Conservation, www.doc.govt.nz/getting-involved/consultations/current/threat-management-plan-review-for-mauis-dolphin/

Dawson SM, Slooten E 1988. Hector's Dolphin Cephalorhynchus hectori: Distribution and abundance. Reports of the International Whaling Commission, Special Issue 9: 315-324.

Dawson SM, Slooten E, DuFresne S, Wade P, Clement D 2004. Small-boat surveys for coastal dolphins: Line-transect surveys for Hector's dolphins *(Cephalorhynchus hectori)*. Fish Bull 201: 441-451

DOC & MFish 2007. Department of Conservation, and Ministry of Agriculture & Fisheries. Hector's dolphin threat management discussion document, April 2007. Available at www.fish.govt.nz/en-nz/Environmental.

Gormley AM, Slooten E, Dawson SM, Barker RJ, Rayment W, du Fresne S, Bräger S 2012. First evidence that marine protected areas can work for marine mammals. Journal of Applied Ecology 49: 474-480.

Hiby L 1999. The objective identification of duplicate sightings in aerial survey for porpoise. In: Marine Mammal Survey and Assessment Methods. Eds. Garner GW, Amstrup SC, Amstrup JL, Laake J, Manly BFJ, McDonald LL, Robertson DG. Balkema, Netherlands, 287 pp.

Laake JL, Borchers DL 2004. Methods for incompolete detection at distance zero. 108-189 in Buckland, S.T., Anderson D.R., Burnham, K.P, Laake, J.L., Borchers, D.L. and Thomas, L. Advanced Distance Sampling. Oxford University Press, Oxford.

MacKenzie DL, Clement DM 2014a. Abundance and distribution of ECSI Hector's dophin. New Zealand Aquatic Environment and Biodiversity Report 123, Report to Ministry for Primary Industries, Wellington, New Zealand. March 2014. www.mpi.govt.nz/document-vault/4350

MacKenzie DL, Clement DM 2014b. Supplementary Material for NZ Aquatic Environment and Biodiversity Report 123S, Ministry for Primary Industries, Wellington, New Zealand. March 2014. www.mpi.govt.nz/document-vault/4351

MPI & DOC 2012. Review of the Maui's dolphin Threat Management Plan. Ministry for Primary Industries and Department of Conservation. ISBN No: 978-0-478-40083-0, ISSN No: 2253-3907, www.doc.govt.nz/ getting-involved/consultations/current/threat-management-plan-review-for-mauis-dolphin/

Rayment W, Dawson SM, Slooten E 2010. Seasonal changes in distribution of Hector's dolphin at Banks Peninsula, New Zealand: implications for protected area design. Aquatic Conservation: Marine and Freshwater Ecosystems 20: 106–116

Rayment W, Dawson SM, Slooten E, Bräger S, DuFresne S, Webster T 2009. Kernel density estimates of alongshore home range of Hector’s dolphins *(Cephalorhynchus hectori)* at Banks Peninsula. Marine Mammal Science 25: 537-556.

Slooten E, Davies N 2011. Hector’s dolphin risk assessments: Old and new analyses show consistent results. Journal of the Royal Society of New Zealand 42: 49-60

Slooten E, Dawson S, DuFresne S 2002. Interactions between Hector’s dolphins and a Golden Bay mussel farm. Report for Marlborough District Council.

Slooten E, Fletcher D and Taylor BL 2000. Accounting for uncertainty in risk assessment: Case study of Hector's dolphin mortality due to gillnet entanglement. Conservation Biology 14: 1264-1270

Slooten E, Lad F 1991. Population biology and conservation of Hector's dolphin. Can J Zool 69: 1701-1707

Thompson FN, Berkenbush K and Abraham ER 2013. Marine mammal bycatch in New Zealand trawl fisheries 1995-96 to 2010-11. New Zealand Aquatic Environment and Biodiversity Report No. 105, Available from Ministry for Primary Industries.

Thomsen F, Ugarte F, Evans PGH 2005. Proceedings of the workshop on Estimation of g(0) in line-transect surveys of cetaceans. European Cetacean Society, 2004. ECS Newsletter No. 44 Special Issue.

Turek J, Slooten E, Dawson S, Rayment W, Turek D 2013. Distribution and abundance of Hector’s dolphins off Otago, New Zealand. New Zealand Journal of Marine and Freshwater Research 47: 181-191

Wilensky U 1999. NetLogo. http://ccl.northwestern.edu/netlogo/ Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, USA.