31 March 2017

BAFFINLAND IRON MINES CORPORATION

Peer Review: Marine Mammal Aerial Surveys in Eclipse Sound, Milne Inlet and Pond Inlet, 1 August - 17 September 2015 (15 March 2016)

Submitted to:
Baffinland Iron Mines Corporation
2275 Upper Middle Road East, Suite 300
Oakville, ON
L6H 0C3

Attention: Megan Lord-Hoyle and Wayne McPhee

REPORT

Report Number: 1663724-002-R-Rev0

Distribution:

1 electronic copy - Baffinland Iron Mines Corporation
1 copy - Golder Associates Ltd.

Table of Contents

FIGURES

1.0 EXECUTIVE SUMMARY

Baffinland Iron Mines Corporation (hereafter, Baffinland) received a Project Certificate for the Mary River Project, Nunavut Territory, on 28 December 2012, and shipped the first ore from a marine port in Milne Inlet in northern Baffin Island in July 2015. Marine transportation of the ore takes place in waters inhabited by marine mammals, predominantly narwhal, bowhead whale, beluga, ringed and bearded seals, walrus, and polar bear. The Terms and Conditions of the Project Certificate require Baffinland to monitor the abundance and distribution of marine mammals, and to study the response of marine mammals, particularly narwhal, to shipping. In 2015, aerial surveys were conducted by LGL Ltd., a consultant engaged by Baffinland, to document the distribution and abundance of narwhal along the shipping route between Milne Inlet, Eclipse Sound and Pond Inlet, and in adjacent areas (Koluktoo Bay and Tremblay Sound). In addition, as 2015 was the first year of operational shipping for the Project, LGL was to assess the potential impacts of shipping traffic on narwhals using an aerial photographic analysis.

Baffinland requested that Golder Associates Ltd. (hereafter, Golder) conduct a third party review of the 2015 marine mammal aerial survey technical report authored by LGL Ltd. (hereafter, LGL), titled *Marine Mammal Aerial Surveys in Eclipse Sound, Milne Inlet and Pond Inlet, 1 August – 17 September 2015* (dated 15 March 2016). Further references to LGL in this report are in reference to this document and not to the company, LGL Ltd. Golder's evaluation was to include, but not necessarily be limited to, the following aspects of the aerial survey technical report: validity of survey design; validity of statistical design; execution of field data collection; execution of data analysis including statistical analysis; interpretation and conclusions drawn from the data; and, review of LGL's recommendations for follow-up studies.

1.1 Extensive Survey

1.1.1 Survey Design

LGL's extensive survey consisted of four surveys flown bi-weekly between 1 August and 17 September 2015, covering transects established in Eclipse Sound, Pond Inlet, Milne Inlet North, Milne Inlet South / Koluktoo Bay, and Tremblay Sound. Surveys were conducted at an altitude of approximately 305 m (1,000 ft) above sea level and a ground speed of ~222 km/h (120 kts). A strip-transect survey methodology was used, with two observers on each side of the aircraft. Observers focused on a ~1,000-m strip on either side of the aircraft, extending from 135.7 m to 1,137.9 m from the midline of the aircraft. Observers also scanned for and recorded sightings beyond the defined strip at a reduced level of effort when no animals were seen close to the aircraft. For each cetacean (whale) sighting, the observer was to dictate into a voice recorder the time, species, number of animals in the group, age class, behaviour, direction of movement relative to aircraft heading, and angle of declination to the aircraft using a clinometer. The clinometer reading would make it possible to calculate the distance of the whale from the aircraft. Observers also recorded sightings of polar bears, seals, walruses and seabirds when possible. The LGL report states that photographs were also collected using a pair of cameras located in the belly of the plane, but there is no subsequent indication that these photographs were used to supplement the analysis of extensive survey data.

Golder identified deficiencies in survey design that could lead to substantial errors in describing the abundance and distribution of marine mammals (the discussion is focused on narwhals since this was the species primarily addressed in the LGL report).

Only one survey per two-week period was conducted in 2015, which is problematic because the distribution and abundance of narwhals in any given area varies a great deal in both space and time. In previous years more than one survey was conducted in each survey period, and in some cases a large difference in numbers of narwhals per stratum per survey period was observed (order of magnitude differences were common, and even differences as large as 80 times as many narwhals in a stratum were observed in two replicated surveys separated by only a day or two). If the surveys were intended to generate a total abundance of narwhals for the whole survey area, this would be of less importance; but when they are used only to present density (number of individuals per km^2) in individual strata (e.g., Milne Inlet North, Milne Inlet South, Tremblay Sound, etc.) of the survey, this can lead to identification of differences between those strata, or between dates, that do not in reality exist. The surveys were used only to present density in individual strata and there was no estimate of overall abundance per survey period.

Another deficiency is the orientation of the survey transects, particularly those located in Milne Inlet North and Tremblay Sound, which ran the length of the water body, rather than perpendicular or zigzag. Transects should be located parallel to any anticipated gradient in abundance. If a survey were to be conducted during a ship passage, and narwhals were expected to move away from the vessel, a transect located immediately over the shipping lane would find very few narwhals, but a transect located off to the side of the shipping lane would find many narwhals. In Tremblay Sound, there was only one long transect, which does not provide any replication and no ability to determine the error in the estimate of narwhal density.

1.1.2 Statistical Design

Narwhal density data were not analysed with a distance-based estimation procedure. Use of clinometer data (to calculate distance) and distance-based estimates (which account for changes in detection rates with distance of the animal from the observer) are essential to obtain more accurate estimates of density or abundance.

Overall, the statistical design was a correlation type of analysis, where no specific treatment effect was controlled in an experimental design. LGL attempted to draw conclusions about the effects of shipping using year-to-year comparisons of the aerial surveys (as these were conducted in years with different numbers of Project vessels travelling through the survey area), but did not account for other obvious disturbances, such as hunting.

Narwhal group size was analysed but was artificially separated into what LGL considered a maximum group size, which would appear to invalidate the observations of group size and its use in statistical analysis.

1.1.3 Field Data Collection

The major issue identified in field data collection was the absence of clinometer data for over half (53%) the narwhal observations.

1.1.4 Data Analysis including Statistical Analysis

The major issue identified in data analysis was related to the decision to use only narwhal data representing a 400-m strip clipped from the original 1,000-m strip in which observations were collected from the aircraft. Without accompanying clinometer data, it is not possible to determine the distance of more than half (53%) of the narwhals in relation to the aircraft, which is required to correctly assign them to inclusion in the 400-m strip or exclusion as being outside the strip. Numbers of animals provided in tables in the report do not appear to sum correctly to either the number that should be assigned to the 400-m strip or the original 1,000-m strip. This suggests that densities were either overestimated or underestimated.

The strip of 400 m was selected based on the citation of a paper by Richards et al. (2010), who identified that detection in their study "dropped by more than half in the first 400 m". Consequently, an uncorrected 400-m strip such as that used by LGL most likely underestimated animal density.

A single model was presented in the results, but there is no indication that the authors ensured it was the best model given the data. The high percentage of zero observations (86%) was stated to have caused difficulties with convergence. It is impossible, given the information provided, to determine whether the model specification fit the data well. Possibly, an alternative model such as a zero-inflated binomial model would have yielded a better outcome.

There was no indication that autocorrelation was investigated, although both temporal and spatial autocorrelation were likely to affect data analysis.

1.1.5 Interpretation and Conclusions Drawn from the Data

The data collected should be sufficient to detect large-scale changes in abundance from year to year, or between survey periods in a given year. However, the subdivision of the data into the spatial 'strata' for which densities were reported is problematic. The spatial scale of the defined strata is much smaller than the likely daily range of narwhals, and earlier years of study found high levels of variation in densities calculated per stratum.

There were inconsistencies in findings identified as significant, such as the stated avoidance of vessels but absence of significant differences between years with high versus low vessel densities, and no accounting for obvious disturbances that have also increased in the study area, such as hunting. The authors do acknowledge the existence of many of these problems, but their importance in potentially invalidating the statistical findings is not sufficiently stressed in the report.

1.1.6 Review of LGL's Recommendations for Follow-up Studies

No specific recommendations were provided.

1.2 Photographic Survey

1.2.1 Survey Design

LGL's photographic surveys were intended to document narwhal occurrence in northern Milne Inlet and Tremblay Sound in three time periods: Before, During, and After a large vessel passage. Specific objectives included documentation of the extent of potential avoidance, the area(s) where narwhals may move to avoid a ship transit, and approximately how long narwhals avoid an area around a vessel.

To minimize potential responses of narwhals to the aircraft, the photographic surveys were conducted at a higher altitude than regular aerial surveys, ~762 m (2,500 ft), and at a ground speed of ~280 km/h (150 kts). One or two observers were on board, but LGL indicated that the data from the observers were used mainly to determine the start and end of the transects for photographic data analysis. Two DSLR cameras located under the belly of the plane, oriented to the sides of the plane at a 27° angle, were used to capture images of the transect at 3-second intervals with a 35.0-mm lens. The images covered an area extending ~1,046 m on either side from the centre line of the aircraft, with a 3 m overlap on centreline.

Transect layout for Milne Inlet and Tremblay Sound in the photographic survey was the same as for the extensive survey, and had the same issues related to orientation and replication of transects.

1.2.2 Statistical Design

Perpendicular distances of each narwhal from the vessel trackline were calculated, but in some cases the narwhals were on the other side of an island from the vessel. The exposure of narwhals to noise would be more appropriately modelled by measuring the shortest distance from the narwhal's location behind the island to the vessel trackline without the island in between.

The photographic analysis did not include a time component that would allow it to address the duration of any avoidance of vessels by narwhals, which was one of the objectives of the study.

The statistical design did not take into account the co-occurrence of other activities that could affect whale distribution, such as hunting.

Issues in the selection of the statistical model (generalized non-linear mixed model), and presentation of data supporting this choice, are as identified for the Extensive Survey.

Narwhal orientation in relation to True North was also analysed. Given that there was variation in vessel tracks both within a voyage and between voyages, and that narwhal orientation was not expressed in a way that related orientation to the vessel, the value of this analysis is very limited. The analysis also did not separate stationary whales from moving ones.

1.2.3 Field Data Collection

No complete sets of Before/During/After narwhal observations allowing comparisons of Milne Inlet and Tremblay Sound were collected in relation to a ship passage. There were no time series in Milne Inlet where clear comparisons of Before vs. During were available.

Some important data for understanding the interaction between narwhals and shipping apparently were not collected. The position of vessels, including both large and small vessels, is essential, as is the collection of data on other potential disturbances, especially hunting.

1.2.4 Data Analysis including Statistical Analysis

The same issues with relation to the generalized non-linear mixed statistical model were present in this analysis as for the extensive survey.

The September 4 photographic survey, which contributed the most data to the statistical model, was conducted when most narwhal were located on the opposite side of Stephens Island from the vessel trackline in the Before period. In addition, a considerable amount of hunting was apparently taking place at the same time. These two factors make any conclusion about narwhal response to large vessels on this date highly questionable.

In addition to the issues about absence of data that relate the orientation of the narwhals to the actual vessel, the requirement for unimodal distribution in the Rayleigh test appears to have been violated.

1.2.5 Interpretation and Conclusions Drawn from the Data

The issues identified above invalidate conclusions drawn from the data.

1.2.6 Review of LGL's Recommendations for Follow-up Studies

No specific recommendations were presented.

2.0 INTRODUCTION

Baffinland Iron Mines Corporation (Baffinland) received a Project Certificate for the Mary River Project on Baffin Island, Nunavut Territory, on 28 December 2012, and the first ore was shipped in July 2015. Ore produced at Baffinland's iron mine at Mary River is transported over land to a marine port at Milne Inlet and transferred to ships, which transport the ore through waters adjacent to Baffin Island. The marine transportation of the ore takes place in waters inhabited by a variety of marine mammals, predominantly bowhead whale, narwhal, beluga, ringed seal, bearded seal, harp seal, walrus and polar bear.

Terms and Conditions attached to the Project Certificate include requirements for monitoring the abundance and distribution of marine mammals. As part of the monitoring program for marine mammals, aerial surveys were carried out to document the distribution and abundance of narwhal along the shipping route between Milne Inlet, Eclipse Sound and Pond Inlet, and adjacent areas (Koluktoo Bay and Tremblay Sound). In order to avoid adverse effects, other Terms and Conditions require studies of the response of marine mammals, particularly narwhal, to shipping. In 2015, the first year of operational shipping for the Project, an aerial photographic analysis of potential impacts of shipping traffic on narwhal was conducted.

Baffinland requested Golder conduct a third party review of the 2015 marine mammal aerial survey technical report authored by LGL Ltd., titled *Marine Mammal Aerial Surveys in Eclipse Sound, Milne Inlet and Pond Inlet, 1 August – 17 September 2015* (dated 15 March 2016). The scope of Golder's evaluation was to include, but not necessarily be limited to, the following aspects of the aerial survey technical report:

- validity of survey design
- validity of statistical design, including use of distance-related data handling and/or density surface mapping
- execution of field data collection
- execution of data analysis including statistical analysis
- interpretation and conclusions drawn from the data
- review of LGL's recommendations for follow-up studies

3.0 GENERAL COMMENTS

The comments in this section provide a high-level summary of major issues and general applicability of the aerial survey program.

There are some issues with the field data collection for the extensive survey, as this report will specify, but the study design and data collected in 2015 for this survey should be sufficient to detect large-scale changes in abundance from year to year or between survey periods within a single year. Nevertheless, a number of methodological concerns were identified and questions remain as to some details of data handling; details are provided in: "Section 4, Specific Comments". The extensive survey would provide better estimates of narwhal density if 'distance' correction were applied in the calculation of narwhal abundance. This method applies a correction factor to account for reduced detection of visually surveyed animals with distance from the observer. Application of the distance correction would require clinometer angles to have been obtained for the majority of narwhal observations (i.e., the ability to determine the distance of the narwhal from the observer), but clinometer angles were collected for only 47% of observations in 2015. It is unclear how presence of whales within the 400-m strip that the authors stated was used in the strip transect methodology was determined without using clinometer readings.

Collection of clinometer data in any future observer-based aerial surveys and use of distance-based estimation is essential if more accurate estimates of abundance are required. However, some aerial surveys conducted for Baffinland in past years have extremely limited clinometer data (e.g., in 2013 and 2014, respectively, 10.3% and 42.5% of observations were collected with accompanying clinometer angles; for pooled 2007, 2008 and 2013 data, only 7.5% of the observations had clinometer angles) (LGL 2015a, 2015b). Consequently, it is not possible to re-analyse the data from year 2015 or earlier years using distance-based estimation. A change to distance-based estimation in future surveys, while improving the accuracy of the density estimate, may introduce issues with comparability of data between years.

Where possible, the study design should be modified so that transects are perpendicular to the ship track across all strata. Orientation of transects parallel to the long axis of the inlet (e.g., northern Milne Inlet) is less likely to provide reliable data than the transverse orientation used in Eclipse Sound/Pond Inlet or the zigzag transects in southern Milne Inlet / Koluktoo Bay. Parallel transect orientation may not be feasible in Milne Inlet, given that the elevation of the terrain near the shoreline in many areas rises abruptly from sea level to several hundred metres. Further, in the case of a photographic survey, the sharp turns and changes in altitude required for the aircraft to conduct a transverse survey in Milne Inlet could cause a large proportion of the photographs to be taken on angles that would make them more difficult to analyse. Zigzag transects similar to those used in southern Milne Inlet / Koluktoo Bay would likely be preferable to parallel transects in northern Milne Inlet for aerial surveys.

Raw data are not presented throughout the report, which makes it difficult for the reader to judge 1) whether the data are distributed well throughout the range of observed values or clumped; 2) whether trends exist throughout the dataset, or whether findings are driven by a few surveys; and 3) whether predictions correctly match observed patterns.

Some important data for understanding the interaction between narwhals and shipping apparently have not been collected, or if collected were not reported. The position of vessels, including both large and small vessels, is essential, as is the collection and interpretation of data on other potential disturbances, especially hunting.

The response of narwhals to large ships along Baffinland's shipping route is inconclusive. Some of the statistical analyses are flawed (see details provided under Specific Comments). Rerunning the analysis with refined models could be an improvement, but this effort should only be completed after data resolution and data gaps are resolved, as these factors would still potentially invalidate or weaken conclusions drawn by improved models. Densities of narwhals are extremely variable both spatially and temporally. Much care is needed before drawing conclusions with correlation types of analysis, where no specific treatment effect has been controlled in an experimental design. There are inconsistencies in apparently significant findings, such as the stated avoidance of vessels but absence of significant differences between years with high versus low vessel densities, and the analysis does not account for obvious disturbances, such as hunting. The authors do acknowledge the existence of many of these problems, but the importance of these problems in potentially invalidating their statistical findings is not sufficiently stressed in the report.

Given the very large spatial and temporal variation in narwhal abundance as well as the difficulty in coordinating the timing of data collection with aerial surveys to provide full Before/During/After coverage to vessel transits, tagging or behavioural studies may be more useful for understanding the response of narwhals to large vessels. Another alternative to examine behavioural effects would be to experimentally change large vessel routes using an appropriate design, to provide more definitive tests of hypothesized effects of vessel transits on whale behaviour.

It was not possible to provide comments on recommendations for further study in the LGL report, as no substantive recommendations were made in the report.

4.0 SPECIFIC COMMENTS: INTRODUCTORY MATERIAL

4.1 Summary

p. 13 line 31: The term “pattern” is ambiguous. The terms used in the analysis should be specified so it is clear what is being addressed.

p. 13 lines 32-34: The following sentence should be restructured for grammatical clarity: “*These results should be interpreted with caution given that there were only two photographic surveys of Milne Inlet with high numbers of narwhals present and during which there were data for all periods Before, During and After a large vessel passage.”* We do agree with the conclusion that the results should be interpreted with caution.

p. 14 line 4-5: The statement “*Narwhal numbers were reduced*” should be replaced with more appropriate terminology that reflects what was measured, such as “*Narwhal observations per unit area in areas in the vicinity of vessel traffic were reduced...”* As written, the language suggests the population has decreased and there is no evidence to support this.

4.2 Introduction (LGL Section 1)

4.2.1 Background (LGL Section 1.1)

No specific comments.

4.2.2 Objectives (LGL Section 1.2)

p. 2 lines 14-31: The following objectives were identified:

1) Document distribution, abundance and movements of narwhals during the open-water season in Eclipse Sound, Milne Inlet and Pond Inlet.

2) Assess the effects of large vessel shipping on the spatial-temporal pattern of narwhal distribution and abundance in and adjacent to Baffinland’s northern shipping route using a statistical modelling approach developed for previous technical reports.

3) Use photographic surveys of narwhals in northern Milne Inlet and Tremblay Sound to document narwhal occurrence Before, During and After a large vessel passage.

    a) determine whether there are movements of narwhals in a substantial portion of their primary summering habitat in response to repeated ship traffic

    b) Document narwhal relative abundance and distribution in relation to a vessel track Before, During and After a transit in order to:

        i) approximate the extent of potential avoidance

        ii) identify the area(s) where narwhals may move to avoid a ship transit

        iii) determine approximately how long narwhals avoid an area around a vessel.

4) Collect and present data on all marine mammal species during the aerial surveys.

p. 2 lines 14-18: Abundance was not assessed in the document, whether on its own or in response to shipping.

p. 2 lines 27-28: The photographic analysis did not include a time component in relation to vessel movements except in a very general sense (i.e., Before/During/After). Therefore, the question of how long narwhals avoid an area around a vessel was not answered.

4.2.3 Narwhals (LGL Section 1.3)

Additional information that should be provided in this section includes the typical and maximum swimming speed of narwhals, and how far they typically swim daily. This would assist in understanding their potential for movement between or within strata in the study area. For example, how long does it take a narwhal to swim from Milne Inlet South to Tremblay Sound, and is this a movement they might make on a daily basis under undisturbed conditions? This speaks to placing the movement of narwhal into context of normal activities and energy expenditures.

4.2.3.1 General Distribution and Population Status (LGL Section 1.3.1)

No specific comments.

4.2.3.2 Migration (LGL Section 1.3.2)

p. 4 lines 14-15: If narwhals begin outmigration by mid-September, why were aerial surveys still taking place on September 17? Surveying during the beginning of the outmigration would likely affect results (density, distribution, movement patterns).

4.2.3.3 Summering Areas (LGL Section 1.3.3)

p. 4 line 37: The sentence “The reasons for narwhal site fidelity to summering areas like Milne Inlet are not clear.” implies that site fidelity is at the spatial scale of Milne Inlet, although the previous paragraph (lines 32-33) indicates “Aerial surveys have also shown that there is much fine scale movement by groups of narwhals among various areas of Milne Inlet and adjacent fjords, both from day to day and over a longer time interval.” The latter statement indicates that site fidelity is not to Milne Inlet as such, and that movement naturally occurs between Milne Inlet, Koluktoo Bay, and Tremblay Sound. This is different from the interpretation of Milne Inlet as a preferred habitat, with narwhals moving to Tremblay Sound only to avoid ships which may be implied by the first statement.

p. 5 lines 3-5: Milne Inlet is identified as a suitable habitat for calving or rearing calves because numerous mother-calf pairs have been observed in the area. What habitat characteristics of Milne Inlet make it suitable for rearing calves? Are there other water bodies in the area with similar habitat characteristics? Does the proportion of mothers with calves vs. the total population of narwhals differ in Milne Inlet versus other water bodies in the study area? As it reads now, there is no mention of other areas that support rearing.

p. 5 lines 7-9: Similarly, Koluktoo Bay has been identified as providing refuge from predators (specifically, killer whales) because Campbell et al. (1988) observed narwhals sheltering from killer whales there. What characteristics make this bay suitable as a refuge, so as to determine what other bodies of water in the area share these characteristics and may also function as refuges from predation? As with rearing, as currently written, there is no indication that other areas nearby can provide refuge.

4.2.3.4 Narwhal Response to Vessels (LGL Section 1.3.4)

p. 5 lines 22-33: This is a confusing description of narwhal response to vessels at Bruce Head. Lowest narwhal counts were recorded when large vessels transited south and highest counts when vessels transited north. According to line 28, narwhals do not respond to northbound transits by leaving the area. Were there observations made during southbound transits? In lines 29-32, in reference to large vessel transits on 18 and 22 August 2015, “many [narwhals] were observed swimming to the south”. It is not specified in this section if these were southbound or northbound vessel transits. However, compilation of data from LGL’s Section 4.3.2 and Table 15 (see compiled table below in comments on LGL Section 3.3.2.1) indicates that 18 and 22 August 2015 were dates when vessels were northbound through Milne Inlet, and narwhal movement to the south seems counterintuitive as an avoidance response. Had they already moved south before the vessel transit? This seems like a strong possibility as it is indicated in the photographic survey that there were virtually no narwhals observed in Milne Inlet on 22 August. Also, there is no mention of other confounding factors (e.g., the presence of shore-based or boat-based hunters was not detected, although there were small boats active in Milne Inlet on both dates).

p.5, lines 26-28: How does this reconcile with the statistical models in this report, which suggest change in movement direction, density, and distance from the track?

p. 5 lines 34-43: How do the behaviours observed in the presence and absence of large vessels compare to literature descriptions of behavioural responses of narwhals to stress (i.e., other studies of response to vessels, or to predators such as killer whales)? Are the behavioural responses to small vs. large vessels similar?

p. 5 line 39: How can the data on change in narwhal swim speed in response to large vessels be “equivocal”? Was there a problem with the data or was the swim speed not measured? Define what is meant by this term.

p. 5, lines 40-43: P=0.049 is marginal. If as stated, the vast majority of the high-count data came from a single vessel transit, the result is suspect.

p. 6 line 3: What are the noise levels of icebreaking vs. large cargo ships? Of small vessels vs. large vessels?

p. 6 line 10: At what distance did the narwhals respond to the oncoming vessels?

p. 6 line 11: How is the “freeze” response (i.e., lying motionless) followed by swimming away from approaching icebreakers, cited here from the literature, different from the behaviour described at Bruce Head on p. 5 line 29 (“groups of narwhals were observed briefly resting while oriented toward the large vessel, before then swimming away and diving”)?

p. 6 line 15: Is it possible to quantify “relatively low levels of noise from the approaching ship”?

p. 6 line 21: If narwhals are disturbed when ships are as much as 130 km distant, it would imply that there is no "Before" analysis possible in 2015. With the amount of vessel traffic that took place not just in Milne Inlet but also in Eclipse Sound and other nearby water bodies, narwhals would be expected to be more or less constantly disturbed in the study area. On the other hand, that 130 km distance is probably based on open ocean sound transmission with no barriers, so the distance of effect may be much lower in the study area considering the more complex land masses.

4.2.3.5 Subsistence Harvest (LGL Section 1.3.5)

p. 6 lines 42-43: The existence of narwhal harvest locations in Koluktoo Bay, Tremblay Sound, Pond Inlet, west shore of Milne Inlet, and near Bruce Head in summer and autumn is very important to the ability to interpret the results of the extensive and photographic surveys due to the potential confounding effects of narwhal response to hunting.

4.2.4 Bowhead Whales (LGL Section 1.4)

4.2.4.1 General Distribution and Population Status (LGL Section 1.4.1)

No specific comments.

4.2.4.2 Migration and Summering Areas (LGL Section 1.4.2)

No specific comments.

4.2.5 Other Marine Mammals (LGL Section 1.5)

No specific comments.

4.3 Description of the Study Area (LGL Section 2)

p. 10, Figure 1: Figure caption should specify which geographic strata were sampled in 2015. The oddly shaped boundary between Milne Inlet North and South may require some explanation.

4.3.1 Ice

No specific comments.

4.3.2 Bathymetry

How are narwhal typically distributed with respect to water depth? Is there a minimum useable water depth? If there is existing literature, it should be mentioned in the preceding section, and the distribution of whale observations in the study area should also be summarized with respect to depth. It has been identified on p. 5 lines 8-9 that narwhals shelter from killer whales in shallow waters; do they respond in the same way to ship traffic?

5.0 SPECIFIC COMMENTS: EXTENSIVE SURVEY METHODS AND RESULTS

5.1 Methods

5.1.1 Survey Design in 2015 (LGL Sections 3.1, 3.1.1 and 3.1.2)

Figure 3: The Milne Inlet North transects run parallel to the shipping track. Depending on the exact location of the middle transect with respect to ship passage, the number of narwhals counted in the transect strip may be very different. For example, if narwhals are repelled from the vessel, then a transect located immediately over the shipping lane would find very few animals, but a transect just off to the side of the shipping lane may find many narwhals, as they move away from the shipping lane. A basic tenet of line-based transect design is that transects should be run parallel to any expected gradient in abundance of the species to be counted (see page 5 in Buckland et al. 2012). As any behavioural displacement disturbance would be expected to generally modify distribution of narwhals along an axis perpendicular to the vessel, establishing transects parallel to the vessel tracks is the opposite of recommended practice and could create significant biases. A spatially weighted selection of perpendicular transects would also be recommended (see Stevens and Olsen 2004), although for logistical reasons, a zig zag pattern would be a reasonable compromise. If 100% coverage was obtained in an aerial survey, the orientation of transects with respect to the vessel track would not be a significant factor. Satellite photography conducted under cloud-free conditions could cover 100% of the study area (eliminating the error potentially introduced by transects) and may provide a better “snapshot” of narwhals over the entire study area.

p. 12 line 6: While modifications to survey design are being discussed, it should be noted in Section 3.1 or 3.1.1 that only one survey was conducted in each two-week survey period in 2015. This differs from the protocol followed in previous years when more than one survey was conducted in each survey period. Given that in other years there was sometimes an order of magnitude difference in the two replicates per stratum per survey period, this difference in survey method may have affected the results, but this potential source of variation is not discussed.

p. 13 lines 10-33. As part of distance analysis, variance is estimated for encounter rate using multiple samples (transects) in a stratum. It is preferred that more than 20 transects are completed within a stratum to adequately estimate this variance (Buckland et al. 2012). With only 4 transects in Milne Inlet, variance estimates associated with encounter rate will likely be unreliable. Since Tremblay Sound has only one transect, there is no measure of variance for its encounter rate estimates. Counterintuitively, the result is that the stratum with the lowest number of transects, Tremblay Sound, will have a lower variance associated with its estimates of density. When only one transect occurs in a stratum, the encounter rate cannot be estimated for the stratum and is only measured for the surveyed area, and so estimates of density should also be limited to the surveyed area.

p. 12 line 17: Given that narwhals are typically migrating away from Milne Inlet by mid-September, results obtained September 17 may not give similar results to those collected during the summer season. Outmigration movement of narwhals may obscure analyses of their numbers or behaviour in relation to shipping activity.

5.1.2 Data Recording Procedures (LGL Section 3.2.1 and 3.2.2)

p. 16 line 28-29. The comparison of photographic data to observer data is important to ensure that estimates from the two survey types produce similar results, as comparisons of the results from the two different survey types are made. Since photographic survey data lack animal movement cues that are available to observers, there is the possibility that photographic survey density estimates are lower, especially during poor weather conditions.

p. 17 line 20-22. The justification of not using full resolution to determine marine mammal sightings, equating using one-ninth of the resolution of the DSLR imagery to human eyesight looking out the window, is not supported with any analysis or citations. What is the basis for this claim? And why not use full resolution to identify the whales versus background noise, or species identification? If the purpose was to use the photos to develop correction factors, wouldn't the best identification possible given the technology be of most use? A discussion of photo analysis methodology with DFO indicated that their procedure is to zoom in to obtain detail in their narwhal photo survey analytical procedure.

5.1.3 Analysis Procedures (LGL Section 3.3.1)

No specific comments.

5.1.3.1 Geographic Strata (LGL Section 3.3.1.1)

p. 19 line 11: It is stated that "survey coverage in these geographic strata varied from year-to-year (Table 2)" but the only details provided in Table 2 are the years in which strata were surveyed, and no information was provided on survey coverage variations within strata. The report should either provide details on coverage within strata (e.g., linear distance surveyed, or percentage of the stratum's area that was covered by the strip transects). Alternatively (if this is what was meant) the sentence should read something like: "Some strata were not surveyed in all years (Table 2)".

5.1.3.2 Density Estimates (LGL Section 3.3.1.2)

p. 19. line 15: Why use uncorrected densities? If photographs had been collected during the extensive survey to supplement the observer counts (i.e., during times that there were too many narwhals to collect clinometer readings), the additional data from the photographs would allow for a more accurate estimate of density including correction for distance. This is an approach that DFO has used successfully (C. Monteith, DFO, pers. comm.). LGL does not appear to have collected photographic data during the extensive surveys in 2015, although photographs were collected concurrently with observer-based surveying in previous years of study (but apparently not analysed, based on data reported in LGL [2015a, b]).

p. 19 line 17: Observations of narwhal were recorded from an observed area that was 1,000 m in width, but the report states that only narwhal counts recorded in a 400-m strip within the 1,000 m observed width were used in the analysis. If no clinometer data were available (as stated in line 15 in the same section), then how were narwhals identified to have occurred in that 400-m strip? It appears that some or all of the data with no clinometer angle were used in density calculations?

p. 19 line 17: Would it not have been possible to use an average clinometer reading for a group of narwhals, given that individuals in a group were often clumped tightly (as stated on p. 15 lines 15-16)? If feasible, this approach has the potential to greatly increase the proportion of narwhals for which distance could have been determined.

p. 19 lines 17-18: A strip width of 400 m was selected based on the citation of Richards et al. 2010. In that reference paper, it was stated that “detection probability dropped by more than half in the first 400 m”, which suggests that an uncorrected 400-m wide strip transect would underestimate animal density. A plot of the extensive-survey narwhal observations, using those collected with clinometer angles, also indicates that the use of this strip width will underestimate density estimates (Figure 1). In addition to illustrating that observations at the far end of the 400-m wide strip were being missed, Figure 1 also indicates that observations close to the aircraft were also being missed within the strip. Overall, this will lead to underestimating the number of narwhal present in the study area.

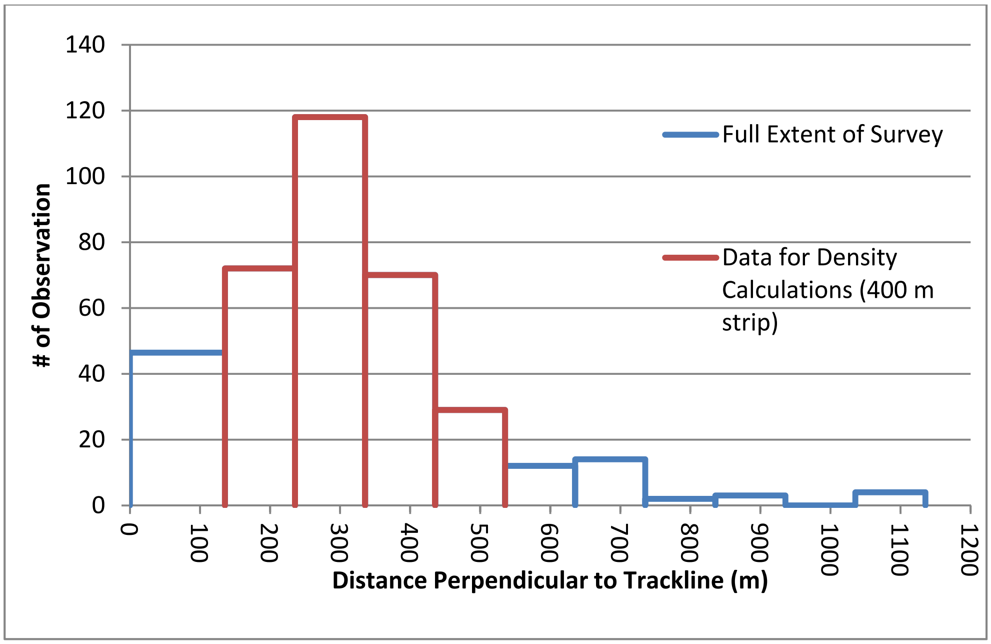

Figure 1: Number of narwhal observations in relation to distance from trackline using the extensive survey data that had recorded clinometer angles. Note that the first bar (0 m to 135.7 m) represents the area under the aircraft that is excluded from analysis

5.1.3.3 Narwhal Group Size (LGL Section 3.3.1.3)

p. 20 lines 11-13: Group sizes in excess of 21 narwhals were subdivided into smaller groups for analysis. Why impose a maximum group size of 21 narwhals (the lower end of the range reported in line 4 of the page) instead of 45 (line 4 of the page)? Also, if previous studies reported sightings of “narwhals with group sizes far greater than these maximum values” (lines 6 to 7), then why not use those “far greater” group size values, or no maximum group size at all? This decision likely invalidates a lot of the group size findings reported later. What were the originally recorded group sizes on photos? In the dataset that Golder received, the observed narwhal group sizes (2013-2015) increased gradually from 1 to 47, followed by a single event where 175 narwhals were recorded. The records of 1 to 47 narwhals could likely be analysed as-is, with the single record of 175 narwhals omitted from this analysis as an outlier.

p. 20 lines 11-13: The aggregation of narwhal counts (line 4) and then disaggregation into what is considered to be a more realistic group size, almost certainly introduces a major source of error into the statistical analyses carried out on these data, which consider number of narwhals per group, and numbers of groups, as response variables. Both of these variables become meaningless after the arbitrary adjustment of group size to which the data were subjected. Rather than disaggregating the data according to an arbitrary procedure it would have been better to leave them aggregated and instead use only the total number of narwhals or narwhal presence/absence as response variables in the statistical models.

p. 20 line 13: In 2015, the dataset received by Golder contained group sizes gradually increasing from 1 to 15, with a single record of 40 narwhals. Was this observation not part of the analyzed data?

5.1.3.4 Shipping Activity (LGL Section 3.3.1.4)

p. 20 lines 28-30: Would it not have been better to consider the timestamp of the AIS data and determine how long before the time of the survey the vessel had passed through a geographic stratum?

5.1.3.5 Calculating Narwhal Closest Points of Approach (Section 3.3.2.2)

p. 21 lines 18-21: It is not clear why distance based on a perpendicular line drawn through land (e.g., an island) from the vessel trackline to a narwhal sighting would have any meaning, as sound will not propagate through the land. The closest distance through the water to the trackline (assuming sound wave propagation in a straight line) would be more meaningful in terms of potential vessel disturbance on whale behaviour?

5.1.3.6 Narwhal Orientation (Section 3.3.2.3)

p. 21 lines 26-27: The effect of the exclusion of data with unknown orientation should be examined to determine if the excluded data were correlated with any of parameters being examined to determine effects (e.g., CPA distances).

p. 21 lines 29-32: The implied assumption in this section is that narwhal orientation is related to movement. Narwhals could be resting, milling, travelling, etc. at the time of the photograph. The resting/milling data do not indicate directed movement away from a vessel, although the milling activity may be a meaningful behavioural response. Were these data excluded from the analysis or were their orientations included in determining the direction of narwhal movement?

5.1.3.7 Density Estimates (Section 3.3.2.4)

It would be preferable for the density estimates to be determined using the same methods as DFO. DFO's 2013 narwhal stock estimate (DFO 2015) report used the following methods:

- The whole survey was recorded using double observers on each side of the aircraft and incorporated into distance analysis – not just a separate analysis and correction factor using a subset.

- Observations with no clinometer angle were used for calculating group size and encounter rate (not applicable to photo data).
- A threshold approach was used for identifying duplicate sightings of whales, with its own coefficient of variation. It could be argued that there are fewer uncertainties when using photos to identify duplicates (although even use of photographs will not reduce the uncertainties to zero).
- Different availability bias correction factors were used for clear versus murky waters.
- The possible relationship between group size, distance and sighting rate was examined (although not found to be significant).

Why was group size not modeled to see if smaller groups were being preferentially missed at greater distances perpendicular to the transect? This likely bias results in an artificially inflated mean group size. Was this process modeled but determined to be not significant?

5.1.4 Statistical Models (LGL Section 3.4.1)

p. 23, lines 18-20: This is a major issue, since a single model was presented in the results, and the authors did not ensure that it was the best model (of those that could converge), given the data. At the least, the variables from the models that were presented in the report could have been dropped, and the reduced models tested against the fuller ones. Surely, the reduced models would have converged, if the fuller models did?

p. 23 line 22: If the high percentage of zeroes was an issue with convergence, why not use zero-inflated Poisson or zero-inflated negative binomial?

5.1.4.1 Response Variable for Extensive Survey Data (LGL Section 3.4.1.1)

No specific comments.

5.1.4.2 Independent Variables and Model Specification (LGL Section 3.4.1.2)

p. 24 lines 64-66: If *NumVess* is calculated for the entire day for the full stratum, then it would make more sense to calculate total within-stratum abundance for each sampling day (using kriging or other spatially-explicit interpolation methods), and use that as the response variable. Since all 2-min intervals within the same day are treated as replicates, degrees of freedom may be artificially inflated.

p. 25 line 7: Why combine Koluktoo, Milne North, and Milne South? While Koluktoo and Milne South had somewhat similar distributions of density over time (high densities at first and third surveys, low densities at second and fourth, as evident from Tables 8-11, and presented in Figure 2), Milne North had a very different distribution, with higher densities at first and second survey (peak at second), and low densities at third and fourth surveys.

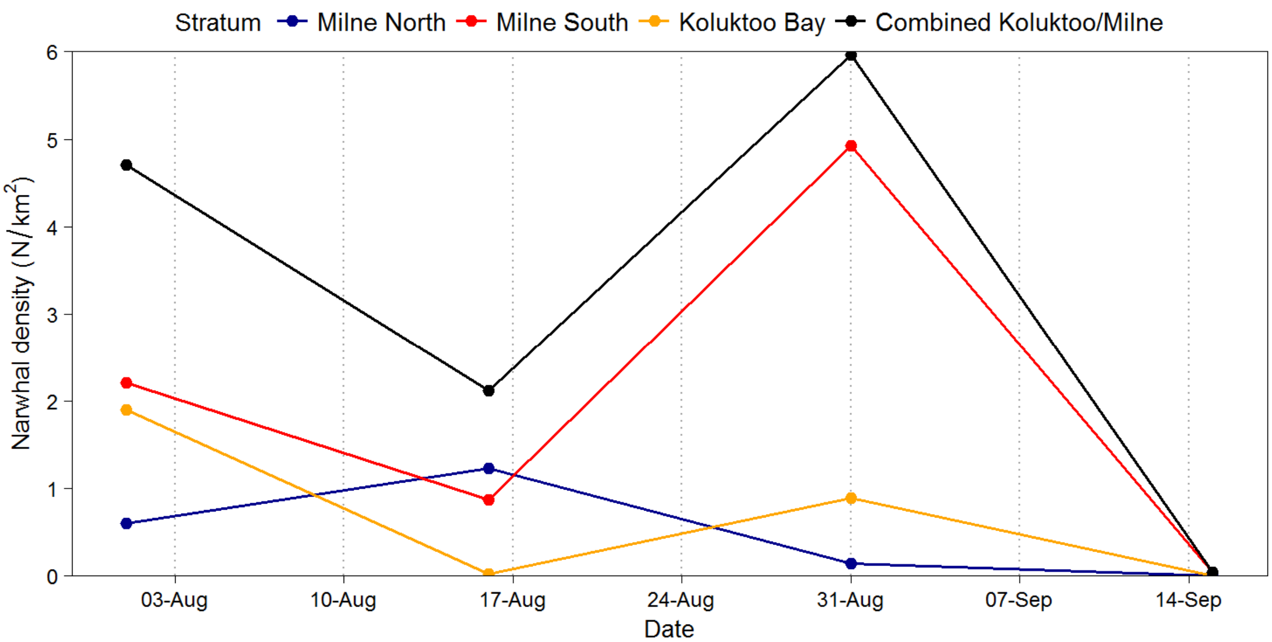

Figure 2: Narwhal density (N/km²), by stratum, based on data provided in LGL's Tables 8-11.

p. 25 equation 2: Why would the model not have an interaction between Julian and Year, and between Year and Stratum? Biologically, it is likely that there were between-year differences in seasonal presence and in relative counts between the strata, respectively.

p. 25 equation 2: Why was Julian not treated as a factor (i.e., Survey1, Survey2, etc.)? Since there were only four surveys in 2015, it might not make sense to represent this variable as continuous. Without seeing the seasonal data as counts/densities time series (not shown at all in the report, which is definitely an issue), it is difficult to know whether it is justifiable to include it as a continuous variable.

p. 25 equation 2: Why was Julian not treated as a random variable in the extensive-survey model, especially considering that Julian was used as a random variable in the photographic model? The objective was not to examine within-season variability in counts, but to characterize the overall abundance/densities.

p. 24 line 56: By Julian date, did the authors mean day of year? Since Julian dates count days starting at noon on January 1, 4713 BC.

p. 24 line 74. It isn't clear why the number of 2-min samples that were available were considered "infinite". The number would be large, however, given a finite amount of time to draw samples in a given day.

5.1.4.3 Testing Differences among Factor Levels and Gauging Effect Size (LGL Section 3.4.1.3)

No specific comments.

5.1.4.4 Model Diagnostics (LGL Section 3.4.1.4)

Major issues

Was autocorrelation examined? There is no mention of it. There can be both spatial and temporal autocorrelation for these data, and neither was mentioned.

5.2 Results

Note that there are some issues with LGL's section numbering in this part of the LGL document.

5.2.1 Survey Effort (LGL Section 4, Tables 3 and 4)

No specific comments.

5.2.2 Survey Conditions (LGL Section 4.2.1)

No specific comments.

5.2.3 Shipping Activity in 2015 (LGL Sections 4.3, 4.3.1)

Major issues

The authors should have included a figure of distribution of vessel counts during each survey – a single bar plot, where each survey date has a panel. Ideally, the figure would also be colour-coded by the different geographic strata used in the model. This would have made the data and the modeling effort considerably more transparent.

The authors should have included a scatterplot of narwhal density as a function of time of year, colour-coded (or paneled) by geographic stratum, and symbol-coded by sampling year. As it stands, it is difficult to compare densities across strata – Tables 9-11 show densities by stratum within each survey, but these are not visually intuitive. Figure 23 compares among-year differences within each sampling season, but does not show the full temporal within-year pattern. Since the within-year pattern is used in the model, it would have been tremendously helpful to include a figure of the time series data.

Minor issues

Section 4.3.1 could be improved by moving all shipping maps and the detailed descriptions to an appendix. Instead, the section should have had detailed plots of the observed densities, preferably summarizing all surveys in a single page, so that it is easier and more intuitive to compare trends within and across strata. Tables of the observed densities can also be moved to an appendix, and just used to provide support for the figures.

5.2.4 Survey Sightings, Density Estimates and Model Results (LGL Section 4.4)

5.2.4.1 Marine Mammal Sightings in 2015 – Overview (LGL Section 4.4.1)

Minor issues

Table 7. The caption says "includes all sightings, including those recorded during connect legs". Does this mean that sightings during connect legs were not included in the models? If so, they should also not be shown in this table (or shown in a separate table), since it is difficult to know what were the actual modeled data.

5.2.4.2 Pinnipeds (LGL Section 4.4.1.1)

There is no discussion of the distribution of unidentified seals (Figure 17) although these far outnumber the identified seals.

5.2.4.3 Narwhals (LGL Section 4.4.2)

5.2.4.3.1 Numbers Observed and Group Size (LGL Section 4.4.2.1)

Major issues

Section 4.4.2.1 and Figure 18. This would have been more useful if the 2015 data were 1) compared to 2007-2014 data; 2) plotted and analyzed as a time series; 3) plotted by geographic stratum; and 4) analyzed in relation to vessel counts, answering questions like "do narwhal groups change with presence of increasing numbers of vessels". As it is, this section is of limited utility.

5.2.4.3.2 General Observations (LGL Section 4.4.2.2)

No specific comments.

5.2.4.3.3 Density Estimates (LGL Section 4.4.2.3)

Major issues

The extensive survey data were to be clipped to the 400-m strip width for density calculations (i.e., all narwhal observations outside of the 400-m strip should have been removed from the analysis), but the data presented in Tables 8, 9, 10 and 11 are not consistent with this process. Based on the provided dataset, there is uncertainty as to what rationale was used to retain observations for density analysis. Regardless of this uncertainty, there are too many observations in the density results tables for periods 3 and 4 (Tables 10 and 11) if observations outside of the 400-m wide strip were dropped from analysis. We do not know what strip width was used for the density calculation; if the 400-m strip width was used with these data, it would indicate an overestimate of animal density, while the use of the original 1,000-m strip width observed from the aircraft would indicate an underestimate.

As previously mentioned, distance cannot be consistently and accurately estimated for observations for which there are no clinometer readings, and therefore it is not possible to consistently and accurately classify the position of these observations with respect to a 400 m subset of the original 1,000-m strip.

Minor issues

Figures 19 through 22: An explanation of “Transect, Acceptable” vs. “Connect, Unacceptable” sightings is needed.

Tables 8 through 11: Columns should be totalled to provide total number of sightings and of individuals for each survey.

Figure 23. As mentioned above, Figure 23 compares among-year differences within each sampling season, but does not show the full temporal within-year pattern. Since the within-year pattern is used in the model, it would have been tremendously helpful to include a time series scatterplot of narwhal density as a function of time of year, colour-coded (or paneled) by geographic strata, and symbol-coded by sampling year.

5.2.4.4 Generalized Non-linear Mixed Model (LGL Section 4.4.3)

Major issues

p. 49 line 29: This text indicates that “86% of data were zeroes”. That seems like a high enough percentage to attempt a zero-inflated negative binomial model. The authors should have included a plot of predicted vs. observed values (as opposed to Figure C-1), or a rootogram (or both), to visually assess the fit of the model throughout the observed counts, but especially in the low-count region. As it stands, it is difficult to know whether the model specification fits the data well.

Figure 24. This figure must also include the raw data. It is very difficult to judge the fit of the model based on the figures/plots presented throughout the report. However, compiling data from Tables 8-11 (see Appendix A, Figure A-1 and Table A-1) and comparison to Figure 24 raises a few questions, and suggests that (as mentioned above) Julian should have been a factor variable, rather than a continuous one:

1) According to Figure 24, Eclipse Sound East has a parabolic change in predicted mean values relative to Julian. However, based on Tables 8-11, the observed densities were 0.0 narwhals/km² throughout all surveys.
2) According to Figure 24, Eclipse Sound West also had a similar parabolic relationship between log-10 narwhal counts and day of year. However, the tables indicate that the observed densities were zero in the first three surveys, and 3.88 narwhals/km² in the last survey, resulting in a non-parabolic relationship with date.
3) The combined Koluktoo Bay/Milne Inlet densities, as derived from Tables 8-11, have a very different relationship with day of year than predicted by the model, with
    a) much higher densities – up to 6 narwhals/km², as opposed to 1 narwhal/km² predicted
    b) a decrease in densities at the second survey, and a strong peak in density at the third survey

p. 49, last paragraph (line numbers are not assigned correctly): Why was *VessNum* not modeled as a continuous variable (or as a dummy + continuous variable)?

p. 49, last paragraph: The term *Year* is in the model as both a main effect and an interaction. Therefore, its significance should not be estimated based on the p-value of the main effect alone. Instead, the term should have been dropped (first from interaction, then from the entire model), and AIC values should have been used to estimate the importance of Year.

p. 49, last paragraph: Since the objective of the model is to find a relationship between the predictive parameters and the observed narwhal densities, variables that were not significant should have been removed from the model (as in the comment above), and if *NumVess* was modeled as a factor, it likely should have also been reparametrized (as 0-2, and 2+) and the new models compared to the original one. In the current version of the analysis, the model is not correctly specified to analyze the collected data.

p. 52, first paragraph. The aggregation of vessel data uses the value >2 in doing the comparison in Table 13. Did the selection of this cutoff for aggregation of multiple vessels affect the results and what were the criteria used for aggregating vessels into groups?

Section 4.4.3: The section should have a table detailing the model's coefficients, so that the fixed parameter effects can be clearly examined.

5.3 Discussion

5.3.1 Narwhal Relative Abundance and General Distribution (LGL Section 5.1)

5.3.1.1 Observed Densities and General Distribution (LGL Section 5.1.1)

No specific comments.

5.3.1.2 Narwhal Movements within the Study Area (LGL Section 5.1.2)

p. 73 line 7: It would be interesting to know how much within-survey period variation there was in the narwhal counts in 2013 and 2014 and how this variation compares with the between-replicate variation that is being examined to quantify effects of vessel traffic in the photographic survey. Note also that the absence of within-survey period replicates in the extensive survey in 2015 limits the utility of comparisons between 2015 and previous years.

p. 73 lines 15-17: If tides and hunting may be important, why were they not discussed? For example, the presence of 19 small boats (of which presumably most if not all are hunting narwhals) along with shore-based hunters during a survey is potentially a major factor in determining narwhal movements.

5.3.1.3 Annual Variation (LGL Section 5.1.3)

p. 73 line 24: The existence of up to 85-fold year to year variability in individual strata suggests that number of narwhals in an area is not a reliable indicator of response to shipping.

5.3.2 Narwhals and Shipping (LGL Sections 5.2, 5.2.1, extensive survey contribution to 5.2.3)

p. 74, last paragraph: If the main and interaction effects were not significant, it is not correct to search for effects in pairwise comparisons. It is difficult to determine here whether the pairwise comparison is valid. The model should have been changed – Year should have been removed, and then tested, to see whether the terms change. It is necessary to see predictions vs. observed values to evaluate the model.

6.0 SPECIFIC COMMENTS: PHOTOGRAPHIC SURVEY

6.1 Methods

6.1.1 Survey Design in 2015 (LGL Sections 3.1, 3.1.2)

Figure 3: As for the extensive survey, the Milne Inlet North transect run parallel to the shipping track. Depending on the exact location of the two middle transects, the results may be strongly different. For example, if narwhals are moving away from the vessel, then a transect located immediately over the shipping lane would find very few animals, but a transect just off to the side of the shipping lane may find many narwhals, as they move away from the shipping lane. A design with zig-zag patterns would be more appropriate for this stratum.

6.1.2 Data Recording Procedures (LGL Section 3.2.2)

p. 16 lines 28-29: Using observer data to verify photo data (rather than vice versa) may be an issue, since observers are more likely to correctly identify animals, especially in less than ideal environmental conditions (e.g., chop or glare situations), where movement is easier to spot than a still picture.

6.1.3 Analysis Procedures (LGL Section 3.3.2)

p. 20 lines 36-37: 28% seems a very high proportion of sightings to be discarded as unsuitable for data analysis. Would this percentage likely be lower at a lower altitude of flight or by using observer data? This should be a consideration for future years of sampling, since discarding a quarter of the data can have a considerable impact on the dataset.

6.1.3.1 Categorizing Data Relative to Vessel Transects (LGL Section 3.3.2.1)

p. 21 lines 9-10: The categorization of the entire transect as “During” (or “Before” or “After”) while analyzing data on the level of a single photograph seems like an issue with replication. If data are analyzed on a photograph level, maybe the variables to use would be time from last vessel and next vessel at the closest coordinates, or distance from vessel, where in the “During” transect it would be an actual value, and in “Before”/”After” it would be multiplied by a “no-vessel” dummy variable. This would have allowed a much higher resolution, while maintaining the photo-level analysis.

p. 21 lines 11-13: The approach described appears to be contradictory. Initially it is stated that the procedure is to select the ship closest to the study area if multiple ships transited the study area in a given day, but then it is indicated that there was never more than one ship transiting the study area during the photographic surveys. The following table combines the information in Section 4.3.2 and Table 15 (p. 55) and clarifies that no more than one ship was transiting the study area at one time. However, on 30 August two ships, and on 4 September three ships, transited the survey area sequentially within the same day. On 4 September, sequential transits were separated by less than one hour between transits. There were no time series on any dates where clear comparisons of Before vs. During were available. Statistical analysis was not carried out for 22 August, for which survey data may have been available for the Before period (not certain if this is the case as the start time of the vessel transit is not identified) because no narwhals were present in Milne Inlet at the time of the survey.

Survey Date	Active Vessels	Vessel Movements	Milne Inlet Surveys (times)	Tremblay Sound Surveys (times)
18 Aug.	Ore carrier M/V Golden Ice	Departed Milne Port, transited northbound through Milne Inlet	Before (1358-1443h) Before/During (1929-2014h)	Surveyed (time not specified) but the photos were not analysed
	6 small boats; no hunting observed	Milne Inlet		
	Tug M/V Svitzer Nerthus	Active in Assomption Harbour		
	Tug M/V Svitzer Njal	Active in Assomption Harbour		
	Sailboat	Active south of Ragged Island		
	Cruise ship M/V Academik Ioffe Ore carrier M/V Golden Saguenay Sailboat Aventura Ore carrier M/V Nordic Olympic	In Eclipse Sound, Pond Inlet, and/or Navy Board Inlet		
22 Aug.	Ore carrier M/V Nordic Odyssey	Transited northbound through Milne Inlet into Eclipse Sound (out of survey area approx. 1315h)	6 replicates surveyed between 0814 and 1456h	During (1306-1316h) After (1501-1512h)
	1 small boat; no hunting observed	Milne Inlet		
	Ore carrier M/V Nordic Olympic	Circling in Eclipse Sound waiting for anchorage at Ragged Island		
30 Aug.	Ore carrier M/V Golden Saguenay	Before survey, southbound in Milne Inlet from Ragged Island During survey, south of Bruce Head, southbound (0833-1125h)		
	Ore carrier M/V Nordic Oshima	Southbound in Milne Inlet to Ragged Island (1340-1458h)	During (1339-1426h) After (1610-1657h) After (1701-1747h)	Before (1326-1334h) During (1435-1446h) After (1755-1805h)
	2 hunting boats	Milne Inlet		

Survey Date	Active Vessels	Vessel Movements	Milne Inlet Surveys (times)	Tremblay Sound Surveys (times)
	Cruise ship M/V Akademik Ioffe Cruise ship M/V Le Soleal Cruise ship M/V Sea Explorer I Ore carrier M/V Nordic Orion	Outside of Milne Inlet		
4 Sept.	Ore carrier M/V Golden Brilliant	Northbound through Milne Inlet into Eclipse Sound (1027-1310h)	Before/During (0952-1038h) During (1041-1123h) During (1127-1210h) During (1214-1300h)	Not surveyed
	Ore carrier M/V Nordic Oshima	Southbound from Ragged Island anchorage through Milne Inlet (1353-1608h)	During (1416-1502h) During (1506-1553h) During/After (1556-1643h)	
	Ore carrier M/V Golden Ruby	Southbound through Milne Inlet to Ragged Island anchorage (1652-1800h)	After Nordic Oshima / During Golden Ruby (1646-1732h)	
	19 sightings of small boats; 1 was close to the narwhal herd for first 4 replicates; 7 sightings of hunting vessels during last 4 replicates; narwhal carcass observed			
	Cargo ship M/V Anna Desgagnes	Westbound from Pond Inlet to southern Navy Board Inlet		

6.1.3.2 Calculating Narwhal Closest Points of Approach (LGL Section 3.3.2.2)

p. 21 line 19: Assuming that narwhals are responding to the noise from ships, the exposure of narwhals to noise may be more appropriately modeled by measuring the shortest distance from the narwhal's location behind the island to the vessel trackline without the island in between, even if that means the connecting line is not perpendicular to the trackline. For example, in the sketch below (Figure 3), the method described in the report would have used line A, but line B may provide a more accurate measure of the narwhal's exposure to ship-generated noise.

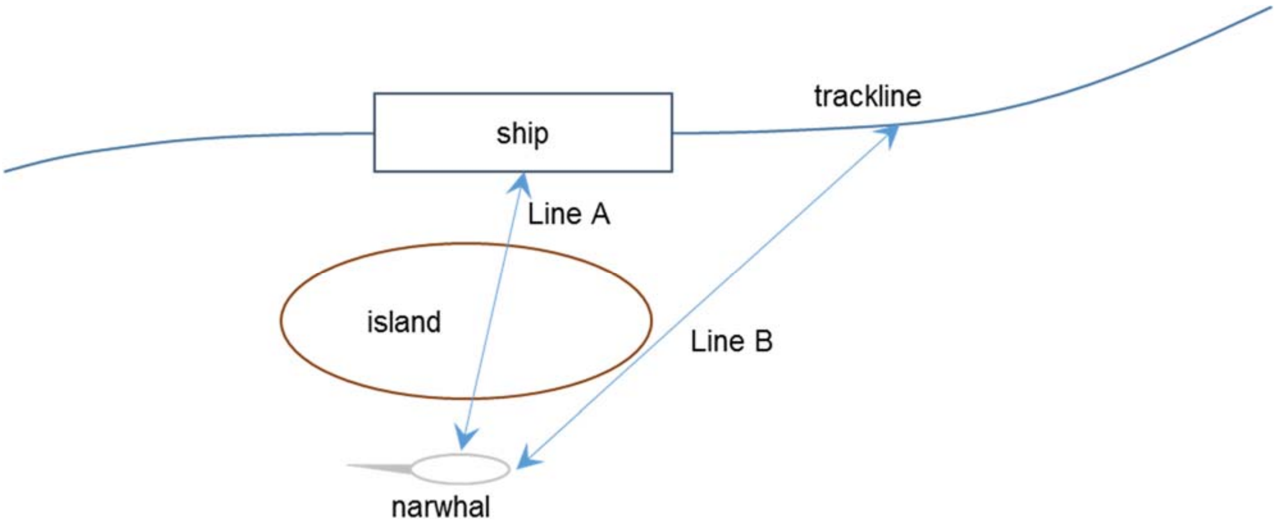

Figure 3: Sketch of estimating distance from narwhal to vessel trackline.

p. 21 line 22: Why calculate CPA from center of photograph, if each sighting's coordinates were calculated, as stated in page 17, line 31?

6.1.3.3 Narwhal Orientation (LGL Section 3.3.2.3)

p. 20, lines 17-18. If tracklines were available (and these were later used to estimate distance from vessel), it should then also be possible to calculate the angle of narwhals relative to the vessels, rather than true north. Ideally, a combination of distance and angle relative to vessel would have been used for the analysis of orientation patterns as function of shipping activity.

p. 21, lines 15-16. Why not also calculate angle to vessel? Instead of analyzing true north angles of the narwhals, analysis of relative angle to vessel would likely be more meaningful in the "During" periods of time.

p. 21 line 24: The description of methods and purpose of this analysis were not clearly explained and it took multiple reads of the material to understand that the analysis was not examining the position of the narwhals with respect to the ship, but the orientation of the narwhal with respect to true North. The value of this analysis is questionable, given that:

1) The orientation of the narwhal is not expressed in relation to the track of the vessel and the nominal vessel track through northern Milne Inlet as shown in Figure 3B changes direction several times; but this is not reflected in the animal's orientation. For example, a narwhal oriented at 90°T is oriented at an angle of 90° to a vessel oriented at 0°T but at 45° to a vessel oriented at 45°T.

2) The direction of the narwhal's orientation is not expressed with respect to the vessel, so it is impossible to determine if the narwhal is facing toward the vessel or away from it.

3) The direction in which narwhals must move in order to avoid vessels changes substantially along the nominal vessel track – for example, narwhals could move to the other side of Ragged Island, to Tremblay Sound, to Eskimo Inlet or Koluktoo Bay, each of which requires a different direction of movement, and the direction of movement would change according to the position of the whale along the vessel track. Thus, pooling all narwhal orientations along the vessel track does not provide a clear understanding of whether directed movements have occurred into an area of refuge.

4) It is unknown whether narwhals are moving or stationary.

6.1.3.4 Density Estimates (LGL Section 3.3.2.4)

No specific comments.

6.1.4 Statistical Models (LGL Section 3.4.1 where methods are carried over to Photographic Surveys, Sections 3.4.2, 3.4.2.1 and 3.4.2.2)

p. 23, lines 18-20: Similar to the extensive surveys, the presentation of a single model in the results is a major issue, as the authors did not ensure that it was the best model (of those that could converge) given the data. At the least, some variables from the model that was presented in the report could have been dropped, and the reduced models tested against the fuller ones. Surely, the reduced models would have converged, if the fuller models did?

p. 23, line 22: Similar to the extensive surveys, if the high percentage of zeroes was an issue with convergence, why did the authors not use zero-inflated Poisson or zero-inflated negative binomial?

p. 24, line 57-63: Even if Water Depth and Distance from Shore were to be excluded from the model specification for the survey data, these are important variables for distribution of narwhal and an analysis of the distribution of whales vs. these variables can add useful information for marine mammal management. This is particularly useful if location of the whales can be determined with good accuracy, as would have been the case for photographic analysis.

p. 26, Section 3.4.1.4: Was autocorrelation examined? There is no mention of it. There can be both spatial and temporal autocorrelation for these data, and neither was mentioned.

p. 27, equation 3. Why no interaction between *BDA* and *LandBarr*? It seems likely that animals would react differently with/without a land barrier between them and the vessel.

p. 27, Section 3.4.2.2: Why was the trackline used to calculate distance between the vessel and the whale, as opposed to the actual position of the vessel, obtained from AIS, at the time when the whale was photographed? The distance between the vessel and the whale, along with orientation of the whale, would have been very informative as to the response of the whale, especially in cases where the whale was only a few hundred metres from the vessel.

p. 20 lines 17-18. In analysis of photographic data, were only the Baffinland vessels used, since they're the only ones that provided trackline data (and if so, why not use the actual position data to determine ship position along the trackline? Or were all AIS data useable for the model in Section 3.4.2.2?

6.2 Results

6.2.1 Survey Effort (LGL Section 4.1.2)

p. 27 lines 74-75: Only a single “Before” survey was used in the analysis. This is a problem, since temporal variability is inherently high for these data, regardless of shipping.

p. 27 line 79: Only one “Before” survey was used for Tremblay Sound as well.

6.2.2 Survey Conditions (LGL Section 4.2.2)

No specific comments.

6.2.3 Shipping Activity (LGL Section 4.3 where specific to Photographic Survey)

The authors should have included a table or a figure of distribution of vessel counts during each transect marked as “During” – a single bar plot, where each survey date has a panel, and the bars depict the number of vessels observed for each “During” transect. Ideally, the figure would also be colour-coded by the different geographic strata used in the model. The figure, or a table, should indicate the number of vessels moving, and number anchored, by stratum, at the time of the photographic surveys.

6.2.4 Photographic Survey Results (LGL Section 4.5)

6.2.4.1 Marine Mammal Sightings – Overview (LGL Section 4.5.1)

No specific comments.

6.2.4.2 Bowhead Whale (LGL Section 4.5.1.1)

No specific comments.

6.2.4.3 Pinnipeds (LGL Section 4.5.1.2)

No specific comments.

6.2.4.4 Narwhals (LGL Section 4.5.2)

6.2.4.4.1 Numbers Observed (LGL Section 4.5.2.1)

The authors should have included a scatterplot of narwhal density as a function of survey date, paneled by stratum, and colour-coded by Before/During/After. As it stands, it is difficult to compare densities across strata – Tables 17-20 provide overall summaries, but it is always preferable to see the full scatter of the raw data. In addition, it is neither intuitive nor easy to track temporal and spatial trends across 4 tables that are spread over several pages.

6.2.4.4.2 Group Size (LGL Section 4.5.2.2)

If the Tremblay Sound data in the Before dataset were not sufficient for χ^2 analysis (presumably due to only observing two groups of three narwhals), why not combine the data into groups of 1, 2, ≥3? This would have allowed a test against the Before data, which is at least as important as testing During vs. After.

Figure 26. This would have been more legible as either dodged bars, or paneled by Before/During/After, rather than stacked bars. The plot should definitely be paneled by stratum, as currently (to our understanding) it displays combined data, which makes it difficult to relate these data to the χ^2 results presented in the section.

6.2.4.4.3 Orientation (LGL Section 4.5.2.3)

This would have been more informative as an analysis of narwhal orientation in relation to the passing vessels in the During transects.

The analysis of narwhal orientation as degrees true north is problematic due to several factors:

1) There is no accounting for distance from vessel in the During transects – the analysis is more likely to find behavioural changes if it accounts for proximity of narwhals to the vessels.

2) The results are compared to vessel passage categorized as north- or southbound. However, the exact angle of each passing vessel relative to True North depends on its location within the shipping lane. For example, in Figure 15, vessels BC4 and BC8 both transited through Milne North, however their angles relative to True North were quite different. Thus, the orientation of the narwhals, which is presented relative to True North, does not represent their orientation with respect to the vessels.

3) The Rayleigh test requires that the distribution be unimodal. Figures 27A, 28A, 28B may indicate violation of this assumption.

4) What is the error bar on Figures 27-28? It was never defined. It is much smaller (~50%) than the reported angular SD values in the text.

5) The breakdown into north- and southbound seems to have resulted in lack of some Before/During/After cells – for both Tremblay Sound and Milne Inlet, 2 combinations of north- or southbound and Before/During/After are missing from the plot. However, these data gaps are not addressed in the results.

Figures 27-38, 30-35: It is very difficult to reconcile the individual survey figures with the combined angular figure. Perhaps if Figures 30-35 were arranged differently, where all “Before” data are presented together (with surveys as panels), and same for “During” and “After”, the information would have been easier to follow.

p. 57, line 13: The six narwhal sightings to the east of the vessel trackline should be removed from the data analysis as the orientation of the narwhal with respect to the ship would be opposite in the east-west direction relative to those on the other side of the ship.

p. 57 lines 32-34: This is very confusing and not particularly meaningful. Analysis of angles relative to the vessel (taking in account distance from vessel) would be more straightforward and intuitive to interpret.

Figures 30-35 are very blurry and hard to read. Likely no need for pictures of the individual vessels.

Table 17: Why are the density values lower than the lower boundary of the confidence interval? Either the means or the confidence intervals are wrong here.

p. 60 line 8. What do the numbers 31, 1, 1 represent? (The reader should not have to track this down in the paper referenced in Appendix B to learn that: 31= sightings seen by both observers, 1 = only seen by observer A, 1 = only seen by Observer B).

p. 60 line 10: Explain why only sightings that were within 200 m of the trackline were used to calculate perception bias. Is this considered the zone of ideal visibility?

p. 60 line 17: Quantify what the authors mean by “relatively low numbers” by providing number per km^2.

p. 61 line 12: The sentence is written in such a way that the reader is expecting four density values representing “both Before and During” for Replicates 2 and 4, respectively. Instead, there is one value for one replicate for Before and another for Before/During. This does not provide any information that can be used to compare density Before to density During.

Section 4.5.2.4 This section could be slimmed by moving all shipping maps and the detailed descriptions to an appendix. Instead, the section should have had detailed plots of the observed densities, preferably summarizing all data in a single page, so that it is easier and more intuitive to compare trends within and across strata. Tables of the observed densities can also be moved to an appendix, as they could then be used simply to provide support for the figures.

6.2.4.5 Generalized Non-linear Mixed Model (LGL Section 4.5.3)

p. 71, Figure 37. The striking conclusion from the comparison of Before, During and After periods is that the observed abundance trends were the same for all, with apparent avoidance of the track line for the first 5-6 km, with densities relatively flat beyond this distance for Before and After, while During the declining trend continues with distance. This suggests that either narwhals have a natural tendency to be distributed declining from the trackline (reflecting bathymetry), or the behaviour of avoiding this area is persistent regardless of vessels in the vicinity. Without some experimental approach where vessels change their tracklines randomly through the region with narwhals present, it is difficult to determine if this is vessel related or an artefact. Also see earlier comments on transects parallel versus perpendicular to the ship tracklines, as this could have potentially created some effects that were not accounted for in the analysis.

p. 71, Figure 37. The raw data should have been presented as a scatterplot overlaid by the predictions, to display the model fit of the data. It is impossible to know whether the predictions are representative of the actual data collected.

6.3 Discussion

6.3.1 Narwhals and Shipping (LGL Section 5.2 specific to Photographic Survey, Section 5.2.2)

p. 75 lines 15-17: Based on acoustic models of the area, which were done for the Environmental Impact Statement, and other acoustic work being done for the client, how far would the vessel noise (at levels that NOAA, for example, considers to be important for marine mammals) be expected to propagate from approaching vessels?

6.3.1.1 Narwhal Orientation (LGL Section 5.2.2.1)

p. 75 lines 20-21: Analysis indicated that narwhals were typically orienting in non-random directions regardless of the presence of a large vessel.

6.3.1.2 Narwhal Group Size (LGL Section 5.2.2.2)

p. 75 line 36: It is important to note that September 4, when narwhal group size doubled, is the date when 19 small vessels, at least some of which were hunting narwhal, were present in the study area, and furthermore a narwhal carcass was observed during the photographic survey.

6.3.1.3 Narwhal Distribution and Relative Abundance (LGL Section 5.2.2.3)

p. 76, line 4: Is the finding that "narwhals were ~1.5 times farther from a large vessel trackline" calculated including the presence of islands, or are narwhals on the far side of an island excluded from this calculation?

p. 76 line 21: The September 4 photographic survey, which contributed the most data to the statistical model, was conducted when most narwhal were located on the opposite side of Stephens Island from the vessel trackline **in the Before period**. In addition, a considerable amount of hunting was apparently taking place at the same time. These two factors make any conclusion about narwhal response to large vessels on this date highly questionable.

p. 76 line 31: Although the report states "It is uncertain how long this response to an ore carrier transit lasted", no evidence has been presented that causally connects the southbound movement of narwhals observed from a shore team near Bruce Head to the southbound transit of the ore carrier Nordic Oshima.

p. 77 line 8: Is this a northbound or southbound transit?

p. 77 lines 28-29: "It is important to account for small boats and hunting behaviour in analyses of narwhal behaviour". We agree, and note that this was not done in the present report.

6.4 Comments on Literature Citations

No specific comments.

6.5 Comments on Appendices

6.5.1 Appendix A: Mapping Narwhal Sightings from Photographs

No specific comments.

6.5.2 Appendix B: Perception Bias Calculations

No specific comments.

6.5.3 Appendix C: GNLMM Model Fit

Figures C-1, C-3. The authors should have included a plot of predicted vs. observed values (as opposed to Figure C-1), or a rootogram (or both), to visually assess the fit of the model throughout the observed counts, but especially in the low-count region. As it stands, it is difficult to know whether the model specification fits the data well.

A plot of residual autocorrelation should have been presented, as data are likely to be autocorrelated.

Figure C-2. It seems that there are some fitting issues based on the residual plot.

6.5.4 Appendix D: Additional Tables

Table D-1. This should have been supplemented by a histogram of frequencies of ship counts within each stratum and year.

7.0 CLOSURE

We trust the information contained in this report is sufficient for your present needs. Should you have any additional questions regarding the project, please do not hesitate to contact us.

GOLDER ASSOCIATES LTD.

Mitch Firman, BSc
Wildlife Biologist

Sima Usvyatsov, PhD
Biological Scientist

Andrea Locke, PhD
Senior Marine Biologist

Dana Schmidt, PhD
Senior Fisheries Biologist

Evan Jones, MASc, PEng, EP(CEA)
Associate, Senior Environmental Specialist

MF/SU/AL/DS/EJ/lmk

\\golder.gds\gal\burnaby\final\2016\3 proj\1663724 baff_marinemammalsurvey_ont\1663724-002-r-rev0\1663724-002-r-rev0-golder peer review marine mammals aerial survey-31mar_17.docx\\golder.gds\gal\burnaby\final\2016\3 proj\1663724 baff_marinemammalsurvey_ont\1663724-002-r-rev0\1663724-002-r-rev0-golder peer review marine mammals aerial survey-31mar_17.docx

8.0 LITERATURE CITED

Buckland S.T., D.R. Anderson, K.P. Burnham, and J.L. Laake. 2012. Distance Sampling: Estimating abundance of biological populations. Springer Netherlands.446 p.

Campbell, R.R., D.B. Yurick, and N.B. Snow. 1988. Predation on narwhals, *Monodon monoceros*, by 19 killer whales, *Orcinus orca*, in the eastern Canadian Arctic. Canadian Field-Naturalist 20 102(4):689–696.

DFO (Fisheries and Oceans Canada). 2015. Abundance estimates of narwhal stocks in the Canadian High Arctic in 2013. Canadian Science Advisory Secretariat Science Advisory Report 2015/46.

LGL Ltd. 2015a. Marine mammal aerial surveys in Eclipse Sound, Milne Inlet, Navy Board Inlet, and Pond Inlet, 31 August – 18 October 2013. Final LGL Report TA-8357-3 for Baffinland Iron Mines Corporation. 6 March 2015.

LGL Ltd. 2015b. Marine mammal aerial surveys in Eclipse Sound, Milne Inlet, Navy Board Inlet, and Pond Inlet, 1 August – 22 October 2014. Final LGL Report FA0024-2 for Baffinland Iron Mines Corporation. 30 November 2015.

Stevens, D. L and A.R. Olsen. 2004. Spatially balanced sampling of natural resources, Journal of the American Statistical Association 99: 262-278.

As a global, employee-owned organisation with over 50 years of experience, Golder Associates is driven by our purpose to engineer earth's development while preserving earth's integrity. We deliver solutions that help our clients achieve their sustainable development goals by providing a wide range of independent consulting, design and construction services in our specialist areas of earth, environment and energy.

For more information, visit golder.com

Africa	+ 27 11 254 4800
Asia	+ 86 21 6258 5522
Australasia	+ 61 3 8862 3500
Europe	+ 44 1628 851851
North America	+ 1 800 275 3281
South America	+ 56 2 2616 2000

solutions@golder.com
www.golder.com

Golder Associates Ltd.
Suite 200 - 2920 Virtual Way
Vancouver, BC, V5M 0C4
Canada
T: +1 (604) 296 4200