ARTICLE
NAMMCO Scientific Publications

ON THE AVAILABILITY BIAS IN NARWHAL ABUNDANCE ESTIMATES

Mads Peter Heide-Jørgensen[1*] & Jochim Lage[2]

[1] *Greenland Institute of Natural Resources, Strandgade 91.2, DK-1401 Copenhagen, Denmark.*
[2] *Jensen Software Sustems, Lammertzweg 19, 24235 Laboe, Germany.*
**Corresponding author: mhj@ghsdk.dk*

ABSTRACT

Abundance estimation of narwhals is usually done with either visual or photographic aerial surveys. The basic estimation for both methods is detection of whales at the surface, and to obtain fully corrected abundance estimates, the at-surface detections need to be corrected for the proportion of whales that, at any given time, is available to be detected at the surface. The surfacing time or 'availability correction factor' is obtained from whales instrumented with dive recorders, that either relay concatenated information on the proportion of time spent at different depth intervals to satellites, or from recovered instruments that collect complete dive profiles, measured at high frequency. Concatenated data binned in depth histograms from Satellite-Linked-Time-Depth-Recorders (SLTDR) falls in two categories, where those that correct the zero depth values with information from the saltwater switch provide larger and apparently more accurate surfacing times, than those collected from instruments that does not correct the zero depth readings. The erroneous detection of near-surface pressure values is likely due to slow response of pressure transducers made from temperature sensitive materials. The high frequency sampling from Acousonde™ recorders documents erroneous surface detections, and adjustments of the dive profiles are needed to obtain realistic near-surface values. Any reconstruction of dive profiles and near-surface values apparently involves some level of correction, and it is recommended, for development of availability correction factors for aerial surveys, that data from zero-adjusted SLTDRs or TDR instruments are used. The mean estimate of surface time from 7 SLTDRs was 29% (CV=0.05). One SLTDR, with steel pressure transducer and zero-adjustments, that was retrieved from the whale, provided a particular long-record (83 days) of reliable high-resolution data. The surface time for this sample was 31%, when calculated as the sum of all depth readings at or above 2 m. The mean of 144 hourly depth readings during 06:00-18:00, and including dives above 3m, was 27.36% (CV=0.8) for 12 days overlapping with the usual timing of aerial surveys. Accurate estimation of smaller depth bins (e.g. 0-1 m) should, even with high resolution instruments, be used with caution when estimating availability bias.

Keywords: Narwhal, aerial surveys, Greenland, diving behavior, surface time, zero offset correction

INTRODUCTION

Estimation of the abundance of Arctic whale populations involves a number of difficulties. Synoptic coverage of large areas of open water is most frequently obtained by visual or photographic aerial surveys (e.g., Richard et al. 2010, Doniol-Valcroze et al. 2020, Watt et al. 2021). The basic estimation is detection of whales at the surface, but a major proportion of the whales are, at any given time, submerged at depths where they cannot be detected on photographs or by visual observers, and in order to estimate the total abundance of whales, it is necessary to correct for the fraction of the population that is 'available' for detection (Marsh & Sinclair 1989). There are currently two methods for estimating the 'availability correction factor', that both depend on instrumentation of Arctic whales. Long-term records (>1 week) of time spent at the surface can be obtained from dive recorders, that relay concatenated information on the proportion of time spent at different depth intervals to satellites (e.g., Watt et al. 2015a). This is usually binned information on the percentage of time within sampling units of hours, that is spent at pre-defined surface intervals (0-1m, 0-2m etc.). High-resolution data can also be collected from retrievable instruments, that sample the entire dive activity of the whales at high frequency (usually at 1 s intervals, e.g. Ngô et al. 2019).

Accurate measurement of the surface, and the depth to which the whales can reliably be detected, is critical for the development of 'availability correction factors. Both types of data collections suffer from either discernible bias in the depth readings at the surface or bias that are unknown and hard to quantify (Heide-Jørgensen & Laidre 2015). Long-term records of dive data may also suffer from sensor drift with increasing bias in surface detection. In addition to bias in surface detection there is also a need for samples that can be considered representative of the average surface time for the population the correction factor is applied to.

There appears to be three sources of error in the estimation of global availability bias. The first source of error stems from the pressure transducer of the instrument used for the recording of dive profiles. The accuracy of the surface detection of the instruments relies entirely on the quality of the pressure transducer involved. Pressure transducers can be inaccurate and cause observation error, especially for the critical detection of the surface (0-2 m depth). Some pressure transducers are slow in their reaction to rapid changes when transiting through water masses with variable temperatures, which is often the case when the whales are approaching the surface. Especially

Heide-Jørgensen, M. P., & Lage, J. (2022). On the availability bias in narwhal abundance estimates. NAMMCO Scientific Publications. *NAMMCO Scientific Publications 12*. https://doi.org/10.7557/3.6518

transducers embedded in heat sensitive material (rubber, plastic, nylon, delrin etc.) will show a delayed response in transducer readings. This will lead to pressure readings that are off for some seconds, with a delayed recognition of the surface. It is possible to correct this error with zero offsetting techniques (zero-offset-correction or ZOC, Luque and Fried 2011), but it is often difficult and somewhat subjective to correct long dive series, where the direction and magnitude of the drift of the pressure transducer varies over time. Metal pressure transducers, as sometimes used in SLTDRs, are less sensitive to temperature changes, have a more rapid response and show more consistent readings of the surface. The optimal solution, to the reduction of the drift of the pressure transducer, is to use software that uses information from the salt-water switch, to calibrate the readings at 0 m depth after each dive.

The second source of error is the duration of the data collection. Long-duration and complete depth profiles are optimal for developing global availability biases, that can be used for correcting aerial surveys, but long duration data series of high resolution are difficult to obtain, as data collection depends on retrieval of archival instruments, that are deployed on the whales for a few days. The third source of error is the number of whales that contribute to the development of the global estimate of availability bias. Ideally sampling should be spread out in proportion to the sex and age groups in the population, and a large sample size is generally preferable, however, this is difficult to achieve, especially over shorter time periods (few years). However, if the surface times can be shown to be consistent between individuals and across age and sex groups, then smaller sample sizes may suffice. When it comes to the time spent at the surface, there are certainly some physiological restrictions that govern how frequent the whales need to be at the surface and for how long they can remain submerged. Almost all studies so far indicate that narwhals, *Monodon monoceros*, for instance, spend between ~21 and ~35 % of their time at the surface (Heide-Jørgensen & Dietz 1995, Heide-Jørgensen et al. 2001, Laidre et al. 2002, Watt et al. 2015a, Westdal et al. 2013). The availability correction determines if the estimate of the abundance at the surface needs to be multiplied by a factor ~3 or 5. It is therefore important to acquire more precise information on the robustness of the surface time estimates. Here data from three different instruments deployed on narwhals (n=33) are analysed, and the shortcomings of each of the datasets collected are presented.

MATERIALS AND METHODS

Satellite-relayed dive data

Sixteen Satellite-Linked-Time-Depth-Recorders (SLTDR) of three types; Mk10 (n=7), SPLASH (n=5) and Mk10 with Stomach-Temperature-Pill (STP, n=4) were deployed on East Greenland narwhals between 2010 and 2014 (Table 1). The Mk10 tags had steel pressure transducers and the SPLASH tags had non-metal pressure transducers.

Table 1: List of 16 SLTDRs that provided data on surfacing time (0-2m) from narwhals in East Greenland. The SPLASH and STP tags were filtered to include data where the sum of all bins was larger than 98 and the surfacing time was within 20 to 35%. For Mk10 tags the filtering includes data during day time hours (06:0018:00) and during summer early fall (1 August through 30 November). The number of bins represent the number of 6-hour periods where the data were collected.

					All tags all seasons		**Mk10 in summer**	
SLTDR	**Year**	**Sex**	**Length**	**Transmitter**	**Number of bins**	**0-2m %**	**0-2m %**	**Number of bins**
20685	2013	M	327	Splash	257	26.61	-	-
20696	2014	M	375	Splash	287	26.25	-	-
21791	2012	M	440	Splash	413	25.12	-	-
21792	2012	F?	280	Splash	286	24.72	-	-
3962	2014	M	414	Splash	291	24.33	-	-
3963	2013	M	400	STP	204	29.35	-	-
3964	2010	M	385	Mk10	78	26.39	27.33	37
3964	2013	M	356	STP	183	27.77	-	-
3965	2013	F	420	STP	218	30.01	-	-
6335	2010	F+calf	395	Mk10	106	26.53	26.69	64
6335	2013	M	390	STP	124	28.30	-	-
7926	2011	M	407	Mk10	68	26.78	31.30	38
93093	2010	M	275	Mk10	37	25.76	27.94	21
93096	2013	F	420	Mk10	12	29.98	31.88	8
93097	2013	F	385	Mk10	83	28.30	31.29	59
93102	2013	M	417	Mk10	54	27.10	27.46	29
Weighted mean	-	-	-	-	-	-	28.88	-
Weighted SD	-	-	-	-	-	-	2.09	-

The tags were programmed to collect and transmit daily data on the fraction of time spent at 0-1 m and 1-2 m depth during four six-hour periods. All the tags had resolution of 0.5 m, they sampled depth every 1 s and dives shallower than 2 m, and dives with durations <20s were ignored. Details of the deployment and programming of these tags are available in Heide-Jørgensen et al. (2015) and Watt et al. (2015b). The Mk10 tags had an automatic correction of the pressure transducer drift using first dry readings, whereas the SPLASH and the Mk10-STP tags did not correct the pressure transducer.

Data from retrievable time-depth-recorders

High resolution dive data (1 Hz) were collected from one whale, #3965, with a depth resolution of 0.5 m (Heide-Jørgensen et al. 2015). The data were obtained from an SLTDR (Mk10-STP) that was retrieved from the whale and contained 83 days of time-depth-recorder (TDR) data when recaptured the following year. The record included dry readings and corrections of the pressure transducer.

High resolution TDR (10 Hz) data were also collected from 17 Acousonde™ tags with non-metal pressure transducers and a depth resolution of 0.5m deployed during August 2013-2019 (see Blackwell et al. 2018, Heide-Jørgensen et al. 2021 for deployment protocol, Table 2).

Table 2: List of narwhals from East Greenland that has contributed data on availability bias during summer (August-September). All whales except #3965 were instrumented with Acousonde™ recorders. #3965 was instrumented with a TDR. Only data from between 06:00 and 18:00 were included and the first 24 hr of the recordings were removed to avoid any effects of tagging. Surface time was calculated with Method 1.

ID	Year	Sex	Length (cm)	Sample (s)	Duration (ds)	Surface time (%)	Surface duration t_s (s) A	Dive duration t_d (s) B	Surface time A/(A+B) %
Asgeir	2018	M	460	302707	7.01	26.67	137	426	24.33
Balder	2016	M	372	283061	6.55	30.25	102	277	26.91
Eistla w/calf	2016	F	~360	100349	2.32	31.20	116	316	26.85
Eske	2017	M	330	34188	0.79	22.79	45	195	18.75
Frederik	2018	M	409	150001	3.47	22.51	80	314	20.30
Freya	2013	F	420	72686	1.68	33.82	172	378	31.27
Frida	2015	F	380	85081	1.97	52.07	89	100	47.09
Helge	2017	M	492	314675	7.28	26.83	128	389	24.76
Helge18	2018	M	492	301178	6.97	20.98	98	413	19.18
Hildur	2017	F	393	316252	7.32	30.04	95	893	9.62
Jonas	2019	M	510	330778	7.66	40.43	183	316	36.67
Kyrri	2018	M	436	432198	10.00	28.39	95	270	26.03
Mutti	2019	F	465	127440	2.95	28.89	115	366	23.91
Nemo	2018	M	410	143805	3.33	21.10	89	388	18.66
Siggi	2018	M	470	129793	3.00	28.19	115	326	26.08
Sigrid	2017	F	379	47750	1.11	26.11	77	250	23.55
Thor	2017	M	457	147581	3.42	28.98	176	474	27.08
#3965	2013	M	420	488302	11.30	33.08	95	216	30.55
Weighted mean	-	-	-	-	-	29.62	112	366	25.51
Weighted SD	-	-	-	-	-	5.99	33	188	7.82

Treatment of dive data

For the SLTDRs it was required that the sum of the time spent in all depth bins should add up to values >98 % to reflect the total allocation of diving time during each 6-hr period. Six-hour periods where the sum was <98 % were discarded, and only data from August through November were included.

The depth recordings from the Acousonde™ tags were down sampled to 1 Hz before a two-step analytical approach was carried out in a modified version of MTDive (https://www.jensen-software.com/mt-dive.html). The first step was to compensate the delayed response of the temperature dependent pressure transducers. When the surface was crossed by the whales this inertia generally results in an 'overshoot' of the depth readings with unrealistic positive depth values.

In order to maintain a constant baseline (i.e. the value of zero assumed to be equal to the water surface), a zero offset correction was deployed. The MTDive function analyze the diving events and a threshold is needed to define the start and end of the dives. The pause between dives is the surface time, hence the criteria for initiation of a dive is important and it was

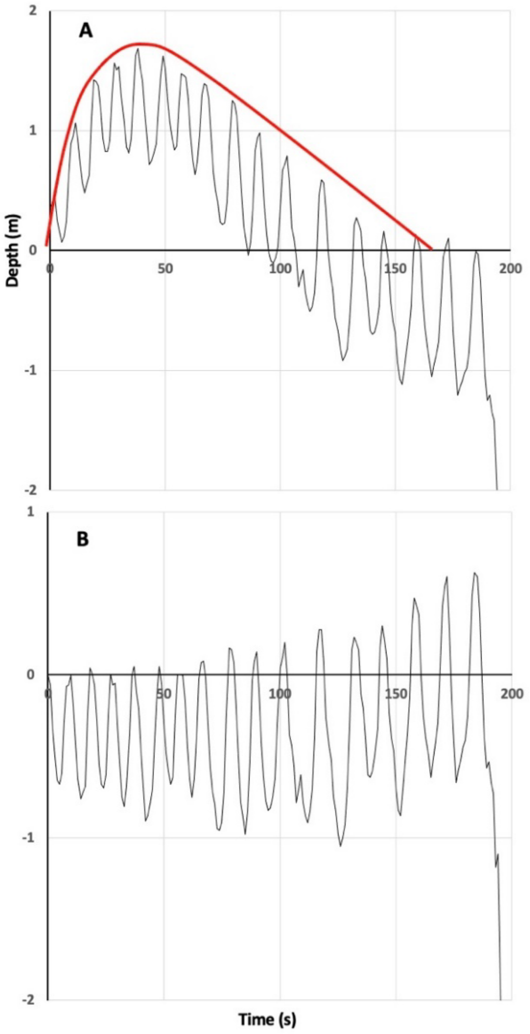

Figure 1: A: Example of a surface event where the inertia of the pressure transducer overshoots the zero m detection of the surface. B: The same event after compensation of the surface detection has been implemented. From Acousonde$^{\mathrm{TM}}$ on Asgeir August 2018.

set to 3 m. The inertia of the sensor and the simultaneous normalization can be described by the exponential function:

$$\mathrm{d}(t) = a \cdot t \cdot e^{(b \cdot t)}$$

with d=depth in meter, t time in seconds in the sampling window (between the two boundary cursors). The parameters a and b are determined by MTDive by curve fitting, and a display of the graph of the function d(t) allows the user to assess the proposed curve adjustment (Figure 1A). If accepted, the measured dive data are subtracted from d(t), and the overshoot is eliminated (Figure 1B). MTDive then searches for the next surface phase and the procedure is repeated to the end of the file. If the proposed curve adjustment is rejected by the user, no subtraction takes place, and MTDive either looks for the next surface event or, by adjusting the window and the zero line, request a new fit. If several submergence phases follow after each other, and the depth sensor is still in decay phase, it is useful to perform the compensation over several submergence phases (Figure 2).

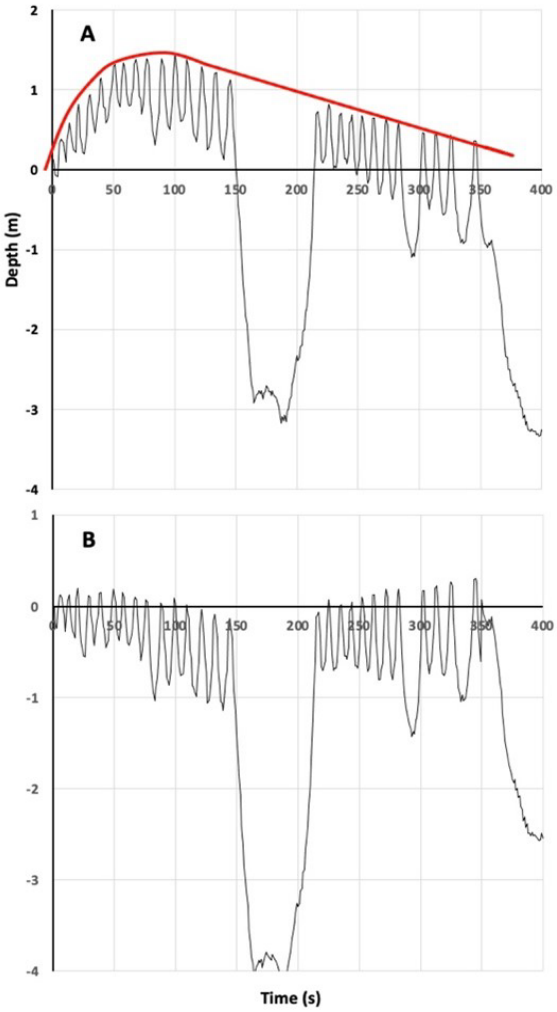

Figure 2: Upper panel: Example of two subsequent surface event where the inertia of the pressure transducer still affects the zero m detection of the surface of the second surface event. Lower panel: The same event after compensation of the surface detection has been implemented. From AcousondeTM on Asgeir August 2018.

The baseline over the entire dive can still vary greatly (Figure 3A) and the elimination of the surface 'overshoot' only affects one surface phase between two dives. If the baseline is low frequency waved then a simple elimination of 'overshoots', which frequently occurs after a dive, by curve fitting is inadequate. Instead, a new baseline is created by elimination of the dive events by interpolation between the surface phases (Figure 3B). This results in a curve which represents a reasonable approximation to the real baseline. In the next step the created baseline is subtracted from the originally measured dive data (Figure 3C) over the entire file of the dive variables, i.e. over the entire duration of the measurements and a zero-offset corrected set of dives is obtained (Figure 3D). The result is a constant baseline over the entire measurement duration, calibrated to the value zero (Figure 3E). Nevertheless, an evaluation of the measurement data with respect to diving depths, diving durations, vertical speeds etc. is unaffected, since MTDive during the analysis adjusts the zero line for each dive individually (Figures 3E and F).

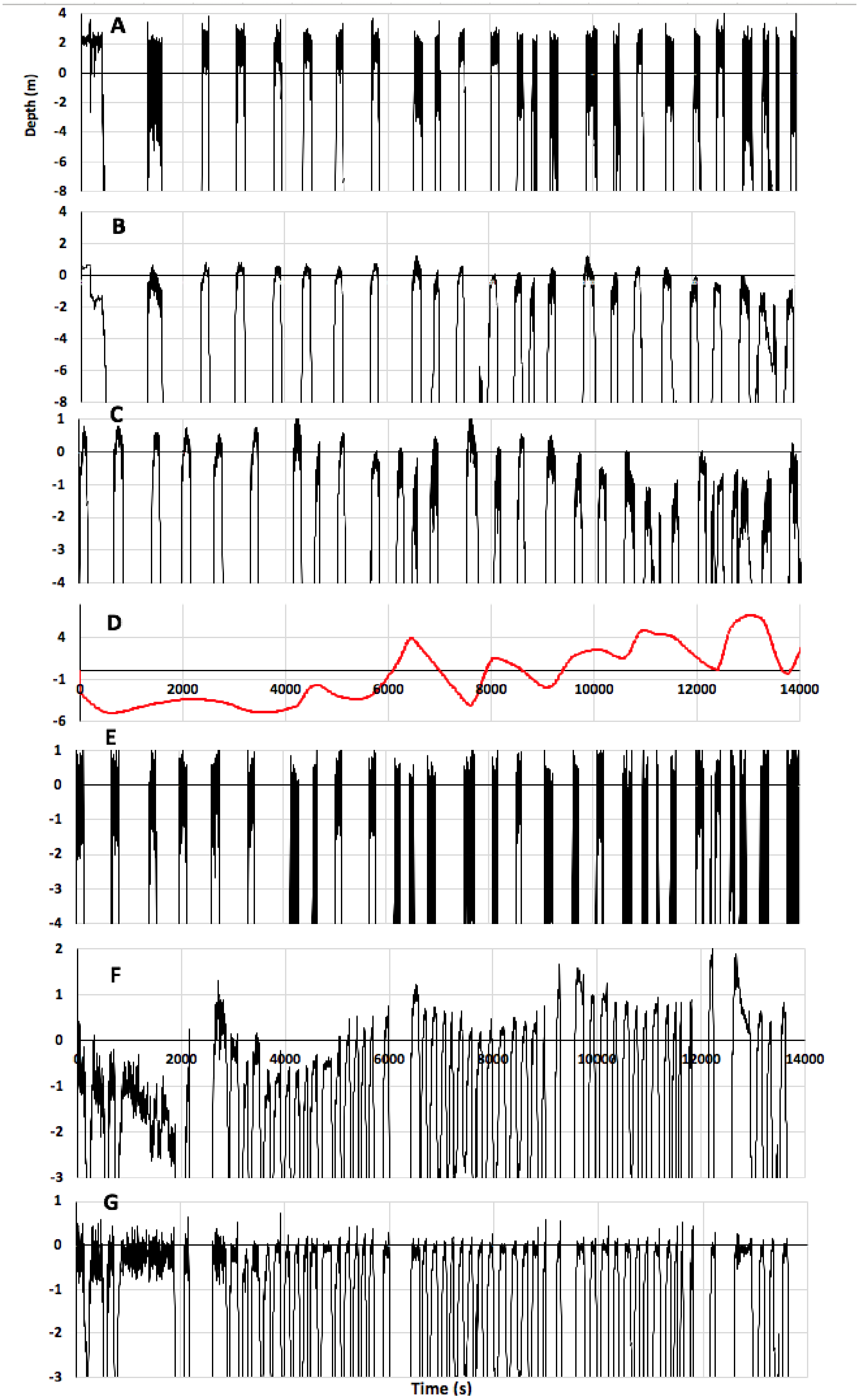

Figure 3: Examples of baseline corrections of the surface (0 m) from Acousonde^{TM} Balder August 2016. A: Uncorrected and corrected baseline (B), C: Uncorrected baseline with overshoot template (D) and zero offset corrected (E). Sample of example of an uncorrected (F) and a zero offset corrected dive series (G)

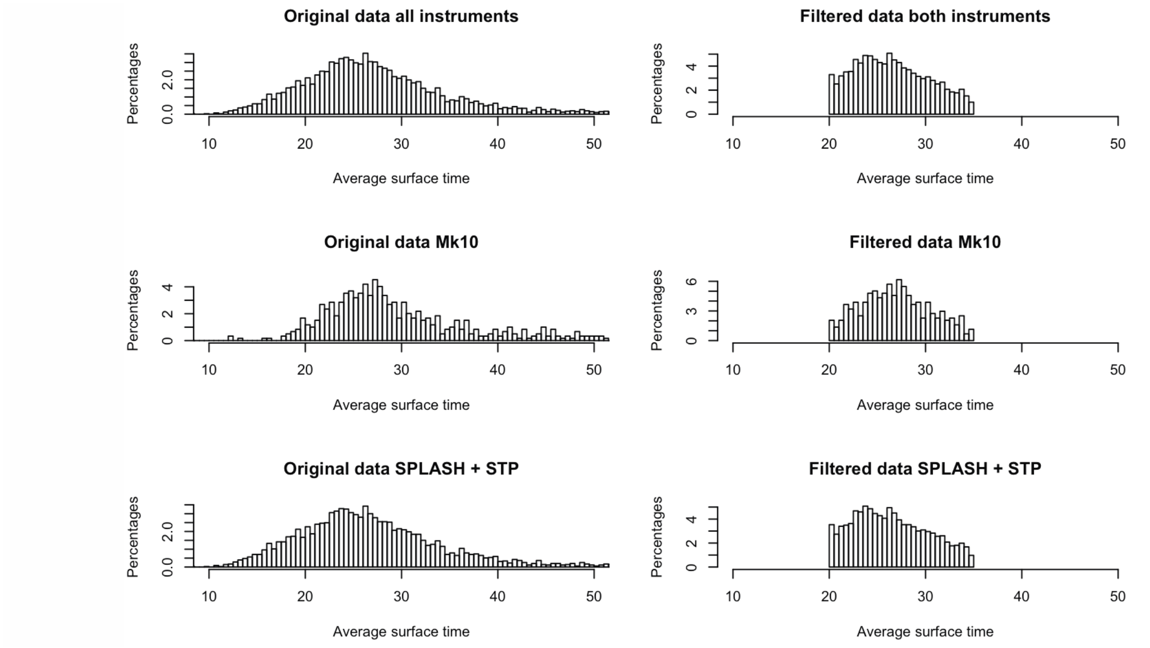

Figure 4: Comparison of the original data on the accumulated time (surfacing time) spent within 0-2m for two transmitter types (Mk10 and Splash/STP tags) before filtering of the data and after filtering. For SPLASH and STP tags the filtering includes data where the sum of all bins was larger than 98 and the surfacing time was within 20 to 35. For Mk10 tags the filtering includes data where the surfacing time was within 20 to 35..

Calculation of surface time

Method 1

Surface time in percentage of the sampling period from 24hr after the whale was released until the end of the record was calculated as:

$$100 \cdot \frac{\sum_{start}^{end} t_s}{\left(\sum_{start}^{end} t_s + \sum_{start}^{end} t_d\right)}$$

where t_s and t_d was the duration of surface (e.g. 0-2 m) and dive periods, and only dives below 3 m were assumed to qualify for a dive to be included.

Method 2

A second method for calculating surface time was:

$$100 \cdot \frac{S}{N}$$

Where *S* is the sum of all seconds at or above the surface threshold (e.g. 0-2m) and *N* is the total sampling period in seconds.

For both methods, only durations of surface periods >1 *s* and <1000 *s* were included in subsequent analyses. Surface periods with >17 min spent at 0-2m are rare and implausible (Tervo et al. 2021) and may be due to sensor failure. Further filtering included daylight hours from 06:00 to 18:00, to match the binned data from the SLTDRs.

RESULTS

SLTDR recordings from East Greenland

A total of 2701 6-hour periods with binned data on narwhal surface time was collected from the 16 SLTDRs (Table 1). The distribution of the surface time showed low and high values outside physiological expectations but with a center around ~25% (Figure 4). When comparing the tags with automatic surface correction (Mk10s) with those without surface correction (SPLASH and STP tags), it is obvious, that those without surface correction generally has lower surface time, and when filtered within a range between 20 and 35%, they still show a left skewed distribution. When filtering the data from the Mk10 and focusing on day time periods during summer and fall the mean surface time is ~29% (Table 1, SD=2.1).

Correction of surface detection on dive recorders

The 17 Acousonde™ tags provided data from 10 hrs to 9 ds but the ability of the recorders to correctly detect the surface varied and generally required both compensation and baseline corrections. Examples of imprecise detection and inertia compensation of the surface from an Acousonde™ tag can be seen in Figures 1 and 2.

A more consistent surface detection can be seen from the recordings from the TDR deployed on *#3965*. It had few surface detections above 0.5 m and there was only a slight drift towards subzero surface (<-0.5m) detections after >5 million depth readings (Figure 5).

When estimating the surface time by hour for the first 12 days of the sampling period from *#3965* (Figure 6), the shortest surface period was 16.25% and the maximum 41.61%. The mean of the 144 hourly readings was 27.36% (SD=4.8) when calculated with Method 2. This is slightly lower than the value of 33.08% calculated with Method 1.

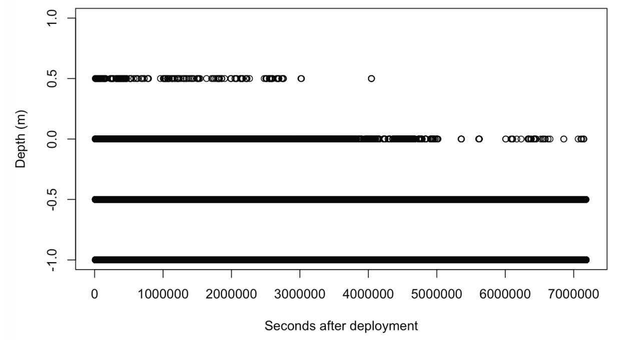

Figure 5: The surface (0.5–2 m) hits from #3965 from >7.000.000 depth recordings. The resolution is 0.5 m.

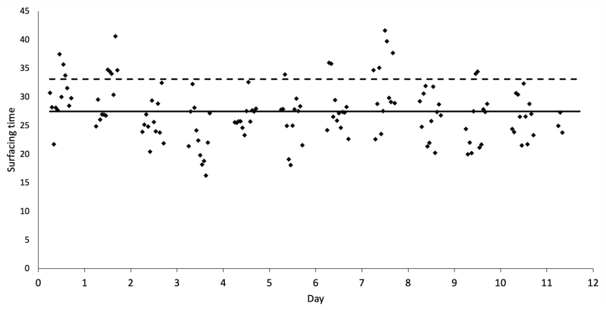

Figure 6: Proportion of time spent at the surface (0-2m) during 12 days and between 6:00 and 18:00 for one whale in East Greenland (#3965) calculated with Method 2. The line shows the mean values from Method 2 (mean=27.4, SD=4.8) and the stippled line indicate mean values from Method 1 (mean=33.1).

Estimates of surface time from dive recorders

Table 2 shows the sampling period and surface time in seconds and the availability bias for 17 whales with Acousonde™ recorders that provided data for more than 24hrs during daylight hours. Two whales had average surface times that exceeded 35% and this might be attributed to short sampling time or bimodal data distributions (Figures 7 and 8). The combined weighted (weighted by the sampling period) average of the proportion of time at the surface for the whales was 29.62 % (SD=6.0). This was affected by the long sampling period of whale *#3965* that had a record long period of 83 days. The average sampling period for the other whales was 4.52 days and the mean proportion of time at the surface for these whales was 29.11% (SD=6.5) which is close to identical to the weighted average for all the whales, but lower than the value of 33.08 % for *#3965* alone (Table 2). The difference is however insignificant (t-test, p>0.5).

The estimate of surface time is highly sensitive to the depth threshold that is used (Table 3). The short events (<2s) with dives below 2 to 2.5. or 3 m will affect the measurements substantially (Figure 9), and if the threshold is increased to 2.5 m the surface time increases with 4 percentage points.

Table 3: Variability in surface time with different depth thresholds for #3965. Only day time periods with measurements within 06:00-18:00 were included. Surface time was calculated with Method 2.

Whale *#3965*	**12ds** 0-2m	**12ds** 0-2.5m	**12ds** 0-3m
Mean (%)	24.91	28.27	31.25
SD	2.80	3.87	4.47

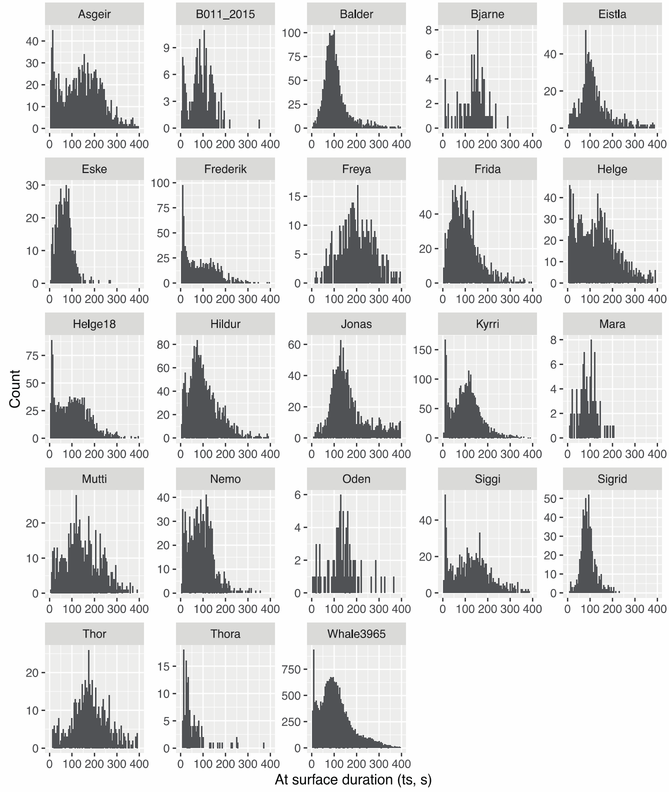

Figure 7: Histograms of the distributions of the duration of time spent at the surface (<2m. t_s) for the 23 whales from East Greenland with high resolution dive records.

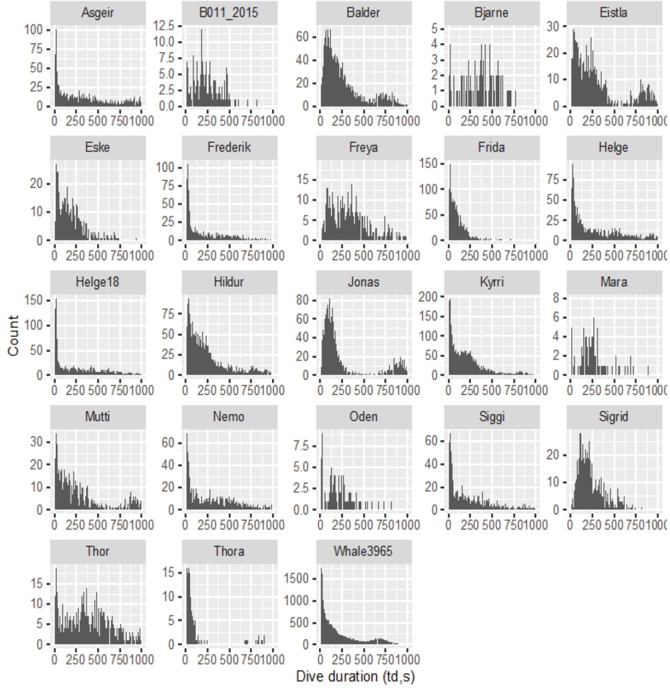

Figure 8: Histograms of the distributions of the duration of time spent on dives below 2 m (t_d) for the 23 whales from East Greenland with high resolution dive records.

Estimates of the duration of dives and duration of time-at-surface

The mean duration of dives (t_d) was 216s for *#3965* and the mean duration of surface periods (t_s) above 2 m depth was 95 s (Table 2). The weighted mean of t_s and t_d for the 17 Acousonde^{TM} tags and *#3965* was 112 s (SD=33) and 366 s (188), respectively. Visual inspection of a subsample of dives from *#3965* show that surface periods of +100 s are common but also that some dives are fluctuating close to the 2 m depth threshold and to avoid unreasonable short dive and surface periods a threshold of 3 m for recognizing a true dive seems necessary (Figure 9).

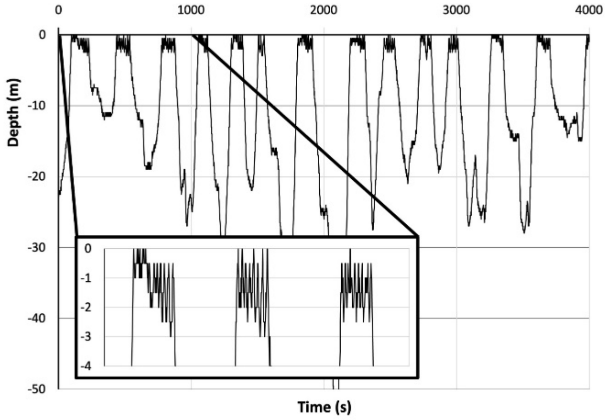

Figure 9: Example of a dive series from #3965 showing the typical +100 s surface duration and +300 s dive durations. The first surface events show depth readings just below 2 m that by definition of a dive (>3 m depth) is included in the duration of the surfacing event (Method 1). The insert show that the 2 m threshold is crossed multiple times leading to unreasonable short dive and surface durations when using the 2 m as a strict threshold.

DISCUSSION AND CONCLUSIONS

Data on diving performance has great implications for the abundance estimates, and the background data for developing availability correction factors needs to be examined carefully and the applied correction factors need to be well documented.

It is obvious that the high-resolution TDR data provide the most reliable information on diving patterns as the full record can be examined and details extracted. Binned data relayed by SLTDRs cannot to the same extent be examined for pressure transducer drift and it is difficult to post factum identify which SLTDRs that provided reliable data. The advantage with SLTDRs is that they can provide data for longer periods (months) and larger samples (more whales), whereas TDR data usually only are collected over short periods of time (days) from few individuals. Among the SLTDRs it is obviously best to use tags that correct the surface detection (0 m) with readings from the salt-water switch and Mk10 (without STP) evidently contributes the most reliable data, whereas data from SPLASH tags should not be used for developing correction factors for availability bias. Heide-Jørgensen & Laidre (2015) identified similar problems with drift of surface readings from tags deployed on humpback whales (*Megaptera novaeangliae*) in 2009-10. They noted that SPLASH tags had a particular large drift in pressure transducer readings, but they also found that Mk10 tags with tag-ware versions 1.24d did not correctly adjust the pressure transducer. Both the Mk10 and the SPLASH tags deployed in this study had later versions of tag-ware.

The high-resolution TDR data suggest an average time spent at the surface of ~30%, which is in good agreement with the duration of the surface periods (t_s) in proportion to the total time for the same whales, when applying a dive criterion of <3m for calculating t_s. A deeper dive criterion (e.g. <10 m) would decrease the time <2 m, because small events that are a relatively large part of the <2 threshold m would be ignored. Both the TDR data from the Acousondes and #3965 show that the t_s values are on average around 100 s and the t_d are above 200 s, but the bimodal distribution of especially the t_d makes it difficult to use simple means. For time series analysis and for surface time estimation as well as estimates of surface and dive durations it seems preferable to adjust the time series of dives for short durations below the threshold to avoid the effect of brief events.

Comparison with data from West Greenland

Heide-Jørgensen & Dietz (1995) presented data for the at-surface-time in the interval 0-5 m for seven narwhals deployed with SLTDRs in Melville Bay in September 1993 and 1994. The pressure transducers used on the tags were of the steel type and are considered to have the same precision as the TDR used on #3965. The average time at surface for the seven whales was 39.3 % (cv=0.13) which is not different for the same interval (0-5 m) for #3965 (36.51%). Previous estimates of surface time of 21% (cv=0.09) and 22% were derived from SLTDRs that were not corrected for drift (Heide-Jørgensen et al. 2010). These estimates should be discarded.

The number of dives per hour for #3965 was consistent around 7.66 dives/h over four months. Narwhals tracked in West Greenland and northern Canada in the 1990s had dive rates that varied between 6 and 9 dives per hour for dives deeper than 8 m with a mean of 7.56 for 7 whales tracked between January and December (Heide-Jørgensen et al. 2001).

Comparison with data from Canada

Watt et al. (2015b) presented data on surface time for 23 narwhals instrumented with SLTDRs in the Canadian Arctic Archipelago, that provided binned and compressed dive data relayed by satellite. Their average surface time was 31.4% in the interval 0-2 m which is similar to what was found by Westdal et al. (2013) for narwhals in Hudson Bay (31.6 %). The precision of depth readings at the surface cannot be assessed from these tags and it is unknown if the pressure transducers were subject to drift.

Apparently, the data set from Watt et al. (2015a) has a higher surface time than the sample (#3965) from East Greenland when estimated with Method 2 but not with Method 1 that ignore dives <3m (Figure 6). This is however not the case for the interval 0-1 m. When compared to the large sample from #3965 from East Greenland, both the samples from SLTDRs deployed in Hudson Bay and in the Canadian Arctic Archipelago are slightly higher. The mean dive rate deeper than 8 m for 4 whales with TDRs in Canada was 8.1 dives per hour (Laidre et al. 2002).

The use of dive cycle information from one whale for abundance estimation

Binned data and point estimates of surface time cannot be used for Hidden Markov Line Transect analysis because this approach depends on complete dive cycle information ideally at a 1 Hz sampling rate and a resolution <0.5 m (Borchers et al. 2013). The average surface time estimates from the seven Mk10 tags, that provided reliable data, are similar to the estimate from one whale (#3965) used for correcting abundance estimates for aerial surveys of narwhals in East and West Greenland (NAMMCO-JCNB Joint Working Group (2021)). The number of dives per hour conducted by #3965 is close to identical to samples of the same dive metrics from narwhals in Canada and West Greenland (Heide-Jørgensen et al. 2001).

For development of availability correction factors the ultimate choice is to use short duration records of variable quality for estimating population means, or to use a few long-term records that better reflects variability across time perhaps with added variance to include population variability. Given that the average surface time is around 26-27% and the average dive rate is about 7.5 dives per hour there is little room for variation in the dive cycle. If accurate surface detection over long duration of deployments is given priority, then the sampling from #3965, with variance based on daily estimates, may best represent the surface time of whales in East Greenland and likely also for other areas.

ADHERENCE TO ANIMAL WELFARE PROTOCOLS

The research presented in this article has been done in accordance with the institutional and national laws and protocols for animal welfare that are applicable in the jurisdictions where the work was conducted.

AUTHOR CONTRIBUTION STATEMENT

MPHJ collected the data and designed the study, JL did the analysis of zero offset corrections of the dive data, and MPHJ wrote the first draft of the manuscript. Both authors contributed to the manuscript revision, read, and approved the submitted version.

ACKNOWLEDGEMENTS

We thank the hunters from Ittoqqortoormiit for their assistance in catching the whales and in deploying and collecting instruments. Outi Tervo, Rikke G. Hansen and Eva Garde are thanked for their help with field logistics, camp maintenance, and capturing operations in Hjørnedal. The study was funded by the Greenland Institute of Natural Resources, the Carlsberg Foundation, Danish Cooperation for the Environment in the Arctic and the Northeast Greenland Environmental Study Program, which is a collaboration between DCE–Danish Centre for Environment and Energy at Aarhus University, the Greenland Institute of Natural Resources, and the Environmental Agency for Mineral Resource Activities of the Government of Greenland.

REFERENCES

Blackwell, S.B., Tervo, O.M., Conrad, A.S., Sinding, M.H.S., Hansen, R.G., Ditlevsen, S. & Heide-Jørgensen, M.P. (2018). Spatial and temporal patterns of sound production in East Greenland narwhals. PLoS ONE 13(6): e0198295. https://doi.org/10.1371/journal.pone.0198295.

Borchers, D.L., Zucchini, W., Heide-Jørgensen, M.P., Cañadas, A., & Langrock, R. (2013). Using hidden Markov models to deal with availability bias on line transect surveys. Biometrics, https://doi.org/10.1111/biom.12049

Doniol-Valcroze, T. Gosselin, J.-F., Pike, D.G., Lawson, J.W., Asselin, N.C., Hedges, K. & Ferguson, S.H. (2020). Distribution and abundance of the Eastern Canada–West Greenland bowhead whale population based on the 2013 High Arctic Cetacean Survey. NAMMCO Scientific Publications 11. https://doi.org/10.7557/3.5315

Heide-Jørgensen, M.P. & Dietz, R. (1995). Some characteristics of narwhal, *Monodon monoceros*, diving behaviour in Baffin Bay. Canadian Journal of Zoology 73, 2106-2119.

Heide-Jørgensen, M.P. & Laidre, K.L. (2015). Surfacing time, availability bias and abundance of humpback whales in West Greenland. Journal of Cetacean Research and Management 15: 1-8.

Heide-Jørgensen, M.P., Hammeken, N., Dietz, R., Orr, J. & Richard, P. (2001). Surfacing Times and Dive Rates for Narwhals (*Monodon monoceros*) and Belugas (*Delphinapterus leucas*). Arctic 54 (3), 284-298.

Heide-Jørgensen, M.P., Laidre, K.L., Burt, M.L., Borchers, D.L., Marques, T.A., Hansen, R.G., Rasmussen M. & Fossette, S. (2010). Abundance of narwhals (*Monodon monoceros* L.) on the hunting grounds in Greenland. Mammalogy 91(5), 1135-1151.

Heide-Jørgensen, M.P., Nielsen, N.H., Hansen, R.G., Schmidt, H.C., Blackwell S.B. & Jørgensen, O.A. (2015). The predictable narwhal: satellite tracking shows behavioural similarities between isolated subpopulations. Journal of Zoology 297, 54-65. https://doi.org/10.1111/jzo.12257

Heide-Jørgensen, M.P., Blackwell, S.B., Tervo, O.M., Samson, A.L., Conrad, A.S., Garde, E., … Ditlevsen, S. (2021). Behavioral response study on seismic airgun and vessel exposures in narwhals. Frontiers in Marine Science. https://doi.org/10.3389/fmars.2021.658173

Laidre, K.L., Heide-Jørgensen, M.P. & Dietz, R. (2002). Diving behaviour of narwhals (*Monodon monoceros*) at two coastal localities in the Canadian Arctic. Canadian Journal of Zoology 80: 624-635.

Luque, S.P., and Fried, R. (2011). Recursive Filtering for Zero Offset Correction of Diving Depth Time Series with GNU R Package diveMove. PLoS ONE 6(1): e15850. https://doi.org/10.1371/journal.pone.0015850.

Marsh, H., & Sinclair, D. (1989). Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. Journal of Wildlife Management, 53(4), 1017–1024. https://doi.org/10.2307/3809604

NAMMCO-JCNB Joint Working Group (2021). Report of the Joint Working Group Meeting of the NAMMCO Scientific Committee Working Group on the Population Status of Narwhal and Beluga in the North Atlantic and the Canada/Greenland Joint Commission on Conservation and Management of Narwhal and Beluga Scientific Working Group. December 2021, Winnipeg, Canada. Retrieved from https://nammco.no/wp-content/uploads/2022/02/report_jwg_2021.pdf

Ngô, C.M., Heide-Jørgensen, M.P. & Ditlevsen S. (2019). Understanding narwhal diving behaviour using Hidden Markov

Models with dependent state distributions and long-range dependence. PloS Computational Biology. https://doi.org/10.1371/journal.pcbi.1006425

Richard, P.R., Laake, J.L., Hobbs, R.C., Heide-Jørgensen, M.P., Asselin, N.C. & Cleator, H. (2010). Baffin Bay narwhal population distribution and numbers: aerial surveys in the Canadian High Arctic, 2002-2004. Arctic 63(1), 85–99.

Tervo, O.M., Ditlevsen, S., Ngô, M.C., Nielsen, N.H., Blackwell, S.B., Williams, T.M. and Heide-Jørgensen, M.P. 2021. Hunting by the stroke: How foraging drives diving behavior and swimming biomechanics of East-Greenland narwhals (*Monodon monoceros*). Frontiers of Marine Science. https://doi.org/10.3389/fmars.2020.596469

Watt, C.A., Marcoux, M., Asselin, N.C., Orr, J.R., & Ferguson, S.H. (2015a). Instantaneous availability bias correction for calculating aerial survey abundance estimates for narwhal (*Monodon monoceros*) in the Canadian High Arctic. DFO Can. Sci. Advis. Sec. Res. Doc.2015/nnn. vi + xx p. Retrieved from https://publications.gc.ca/collections/collection_2015/mpo-dfo/Fs70-5-2015-044-eng.pdf

Watt, C.A., Orr, J.R., Heide-Jørgensen, M.P., Nielsen, N.H. & Ferguson, N.H. (2015b). Differences in dive behavior among the world's three narwhal (*Monodon monoceros*) populations correspond with dietary differences. Marine Ecology Progress Series 525, 273–285. https://doi.org/ 10.3354/meps11202

Watt, C.A., Marcoux, M., Ferguson, S.H., Hammill, M.O. & Matthews, C.J.D. (2021). Population dynamics of the threatened Cumberland Sound beluga (*Delphinapterus leucas*) population. Arctic Science. 7(2): 545-566. https://doi.org/10.1139/as-2019-0030

Westdal, K.H., Richard, P.R. & Orr, J.R. (2013). Availability bias in population survey of Northern Hudson Bay narwhal (*Monodon monoceros*). Polar Biology 3, 1235–1241.