CALIBRATING LARSEN-500 LIDAR BATHYMETRY IN DOLPHIN AND UNION STRAIT USING DENSE ACOUSTIC GROUND-TRUTH

by Rob HARE¹

•

Abstract

Dolphin and Union Strait, in the Canadian Arctic, was surveyed in the summer of 1990 using the Larsen 500 scanning Lidar bathymeter. As part of the 1993 Arctic hydrographic surveys, five test areas within the Lidar-surveyed area were chosen for detailed acoustic measurements. These data were used to calibrate the Lidar depths prior to integration with acoustic soundings in hydrographic field sheet construction.

For the calibration, geostatistical software was used to bring the two data types to common grid points, and to estimate the total error surface of both Lidar and acoustic measurements due to errors in depth measurement and reduction, positioning system errors and errors in interpolation due to the roughness of the bathymetry. The differences in the two data sets are examined and the reasons for biases are discussed. A precision estimate of the Lidar depth measurements was made based on the difference residuals and the *a priori* estimates of precision of each of the measured data sets.

1. INTRODUCTION

Charting the Canadian Arctic has always been a difficult and expensive proposition. The short ice-free season, with longer daylight hours has been used to good advantage by ship-borne hydrographic surveys for many years. By using an aircraft as the surveying platform both the mobilization and data collection speed can be greatly increased. The areal coverage of bathymetry in Arctic waters, and hence the reliability of the charts, is arguably as important as the accuracy, since much of the area is only sparsely surveyed.

¹ Canadian Hydrographic Service (CHS) - Pacific Region, Sidney, British Columbia, Canada.

Due to the sounding grid spacing that is typical of Lidar surveys (about 30 metres), additional tools may be needed to generate statistical estimates of the confidence in the bathymetric surface generated by the somewhat sparser data. This paper examines the precision and accuracy of Lidar surveys as carried out in the Canadian Arctic and discusses the advantages and disadvantages of the method used to calibrate the Lidar measurements.

2. BACKGROUND

A consortium of mining companies is looking to exploit poly-metallic sulphides in the Coppermine area of the Northwest Territories (N.W.T.), in Northern Canada. Ice-breaking bulk carriers, with 12 metres static draught, will be taking the sulphides to Japanese or European markets via Dolphin and Union Strait to the West or Coronation Gulf to the East (Fig. 1). The route to the East passes through the less surveyed Victoria Strait, which has a history of heavy ice conditions and is known to be shallow and hazardous. A more northern route through Parry Channel, which is deep and has good ice conditions, is accessed from Prince of Wales Strait, which runs northerly from Amundsen Gulf at the northwest end of Dolphin and Union Strait. Corridors through the two main channels; Lambert Channel to the West and Cache Point Channel to the East (Fig. 2) were surveyed in detail using acoustic methods. Lidar measurements, using the Larsen-500 system, were also made in the areas where the water depth did not exceed about 30 metres (the maximum depth is limited by the turbidity of the water).

In order to calibrate the Lidar measurements, five, one km² test areas in Dolphin and Union Strait (labeled A-E in Figure 2) representing a range of depths were selected for dense acoustic coverage. Additional acoustic measurements were made in order to fill in gaps between Lidar lines due to lost data and to collect depths in areas where the water depth exceeded 30 metres. Two corridors were surveyed acoustically, one in Cache Point Channel and the other in Lambert Channel (not shown), and shoal examinations were performed. An examination of the reliability of the corrected Lidar depths will be performed by comparing them to the acoustically determined shoals and the surveyed corridors. The results of this comparison will be presented in a subsequent paper.

3. POTENTIAL ERROR SOURCES

Scattering and Undercutting

The true depth (and position offset) of each sounding could be simply calculated by Snell's Law if the Lidar pulse followed a straight-line path (Fig. 3a). However, because of the scattering of photons in the water column, due to the turbidity of the water, a bias can be introduced in the Lidar measured depths which may be either too shallow for off-vertical beams when the refraction causes the ray path to be bent towards the vertical (3c), or too deep when the ray path is erratic



FIG. 1.- Area Location Diagram.

due to scattering (3b). The deep-biasing problem is more typical of Lidar scan angles close to the nadir and the shoal-biasing is more common at angles of greater than 15 degrees from the nadir. The Larsen 500 system uses a near-constant scan angle of 15 degrees by firing Lidar pulses in a conical pattern. Biases introduced as a result of water turbidity are minimized at incident angles near 15 degrees [GUENTHER et. al. 1984]. Although the figure implies there is a refraction correction to be made, in fact this is never done. The scattering fans the signal out so that it swallows up any refraction. The mean photon path is nearly always close to vertical.

Beamwidth

Further errors in the depth can result from the spreading of the Lidar beam in the water column. The illuminated area - the so-called "footprint" - on the seafloor grows with depth much in the same way as the insonified area from an echo sounder. This may cause one of two phenomena:

1. The shallowest part of the seafloor within the footprint may cause enough "early-returning energy" in the peak to be detected, or



FIG. 2.- Portion of CHS chart 7776 (Dolphin and Union Strait) showing five test areas and acoustically-surveyed corridor through Cache Point Channel. Soundings have been removed for clarity.

 An averaging of depths within the footprint may take place in order for there to be enough return energy to be detected - especially in deeper water or when the water is very turbid, causing low signal to noise ratio.

These phenomena are not uncommon to echo-sounders and although the beamwidth is narrower (in our case the width between the half power points of the acoustic beam is 9 degrees and for the Lidar about 28 degrees). The Lidar spot is already 2 metres in diameter at the water surface whereas the echo sounder pulse is just being introduced into the water at the transducer (effectively a point energy source). The spreading characteristics are demonstrated in Figure 4.



FIG. 3.- Lidar biases introduced by turbid waters (e.g. due to phytoplankton).

			500 metres				Ũ
Acoustic						Ŷ	LIDAR
0 m		1	Surface				<u>rootprint</u> 2 m
0.5 m			2.9 metres (Area "C")	1	Ι		3.4 m
1.1 m			7.2 metres (Area "D")				5.5 m
2.9 m			18.6 metres (Area "A")	/			11 m
3.6 m	Γ		22.7 metres (Area "B")				13 m
4.4 m			28 metres (Area "E")				15.5 m

FIG. 4.- Acoustic and Lidar footprint sizes for average depths of test areas.

Temporal Change

Another possible error source is due to the temporal aspect of the seafloor. The Lidar bathymetry was collected in 1990 and the acoustic ground-truth in 1993. Because of the predominance of ice in the Canadian Arctic, movement of the seafloor, due to ice scouring in the channels and ramping of ice in shallow waters close to shore, is not unexpected. Following the line of thinking of Velberg [1993], we can say that the mean seafloor depth over a large enough region will stay more or less the same, since the amount of seafloor materials has not changed except for a minimal amount of siltation in the three years between surveys. Any change in the shape of the seafloor will however increase the amount of noise in the measurement of differences between the two sets of data.

Bottom type

The nature of the seafloor plays an important role in determining if what the two systems "saw" is the same. The test areas were not sampled for bottom type but samples taken by grab in other parts of Dolphin and Union Strait show that the bottom is generally a mix of hard sand, shell and coral. The strength of the return signal is determined by the hardness of the seafloor in the context of acoustic measurements and by the brightness in terms of the Lidar measurements. An analogy would be that a rock bottom is to echo-sounders as a white bottom is to Lidar.

Predicted versus Observed Tides

Further errors may be introduced due to differences in the tides used for the two surveys. The gauge installed at Bernard Harbour in 1990, some 20 miles to the northwest of the test areas, ceased operating part way through the fourth day of the Lidar survey (Julian day 230). As a result, predicted tides, adjusted for pressure readings taken at Coppermine were used to reduce the Lidar soundings. Three tidal zones were extrapolated from the Bernard Harbour gauge - area E falling into zone 2 and areas A-D falling in the extrapolated Camping Island zone.

The tides for the 1993 survey are based on observed tides from a gauge installed on Camping Island, which is in the same tidal zone as areas A-D. Area E tides were interpolated using observed tides from Camping and Bernard Harbour gauges. The complexity of the tidal signal in the area makes interpolation of zone boundaries difficult, since the signal changes from being mainly semi-diurnal at the north end of the Strait to diurnal at the south end of the Strait. The 1993 observed data was used to determine the location of the zone boundaries used for the 1990 extrapolation.

The expected errors in the 1993 tides are on the order of a few centimetres, at least for areas A-D. The expected errors in the 1990 predicted tides could exceed 0.1 metres, perhaps for a period of several hours - long enough for the entire test area to be covered by Lidar. This error would show up as a bias in all depths for an entire flight line. Note that the time taken to survey the five areas using acoustic measurements was 37 hours, 38 minutes while the Larsen-500 surveyed all of the approximately four lines required for each test area in 07 hours, 15 minutes. There is some sacrifice in data density for such rapid coverage as will be shown in the next section.

96

Signal Processing Techniques

The last identified error source is due to the superposition of the signal reflected from the seafloor with that of the back-scatter from the water column [e.g. WONG et al., 1993]. When the water is very clear, the water column back-scatter is very small and so reflections from deeper areas have little interference (Fig. 5b), but in shallower waters the superposition of the two return signals can cause the seafloor return to be displaced to the left (earlier in time) thus causing the depth estimate to be too shallow (Fig. 5a). A marginal gain in accuracy might be achieved in shallow waters by more sophisticated signal processing techniques (e.g. waveform decomposition), but at a cost of increased processing time. Ground-truthing of these algorithms is still needed.



FIG. 5.- Superposition of return signals from sea-surface and seafloor.

4. PROJECT APPROACH

Comparisons of Lidar data to acoustic ground truth have been made in the past by visually examining overlapping sources at the same scale and extracting differences where soundings from the two sources are nearly coincident. Another method is to map one data set onto a Digital Terrain Model (DTM) built from the other data set and use the differences for analysis of noise and biases. Both of these methods ignore the problems of the roughness component of the seafloor and make the incorrect assumption that both data sets are without position error. This of course is never the case, although precise GPS is making this problem less of a concern.

The approach taken in this paper is to use a Canadian geostatistical software package, HYDROSTAT [KIELLAND and DAGBERT, 1992], developed for the CHS by Geostat Systems International Inc. of Laval, Québec, Canada, to bring each of the data sets to common points on a regular grid. HYDROSTAT is the geostatistical portion of IHOstat, the proposed IHO-accepted data quality assessment software. Because this software uses the Kriging method, error estimates are available as a biproduct of the gridding process. These error estimates are based on the weighting function which uses the variograms computed from the acoustic profile data to determine the roughness of the seafloor. The software also incorporates the position error estimates and the errors in the reduced depths into the total error estimate for each grid point.

Because of the size of the Lidar footprint (recall Fig. 4) and the relative sparseness of the data (30 metre spacing) the acoustic profiles with 2 to 3 metre along-track spacing were used exclusively in the computation of the variograms. These variograms, containing the information about the roughness of the seafloor, were used to krige both the acoustic and the Lidar data. Any averaging of the seafloor roughness as a result of the size of the footprint will naturally tend to reduce the slope of the variogram and make for optimistic error estimates.

5. DATA ANALYSIS

Both data sets were reduced for tides: the 1990 Lidar surveys using predicted tides and the 1993 acoustic surveys using observed tides. All other reductions were performed (variable launch draft, mean velocity, etc.) on the acoustic data and an estimate of the contribution of each reduction to the total error of the reduced soundings was estimated in an error budget, much in the same way as proposed by MYRES [1990]. An estimate of the error contribution from the sea state conditions was also made, from an examination of the original echo-grams. The quadratic sum of all components of error is given for one of the data sets in Table 1. A summary of the sounding error estimates for the acoustic data sets given in Table 2.

Description	Comments	Precision (m)	Shoal Bias
Sound Velocity measurement	0.02 (%)	0.002	0
Temporal variation of velocity	0.5 (m/s)	0.003	0
Application of SV	using depth-weighted mean	0.010	-0.02
Sounder accuracy (m)	function of range, depth and resolution	0.166	0
Heave	p-p estimated at 0.2 metres	0.100	0
Draft measurement error	Variability in launch loading	0.050	0
Tidal Measurement	due to distance from gauge	0.100	0
Draft application	Due to step function	0.050	0
Errors due to Speed over ground	Algorithm uses GPS velocity	0.020	0
Application of tide	Due to linear interpolation and timing	0.005	0
RSS depth precision:		0.230	-0.02

Table 1 - Example sounding error assessment at 90%(the shoal bias has not been applied).

Test Area	Est. (90%) depth error	Number of soundings	Number of variograms	
А	0.213	32603	2886	
В	0.214	33022	4189	
С	0.212	50958	2791	
D	0.230	30277	4439	
E	0.256	18708	1639	

Table 2 - Summary information on acoustic ground-truth areas.

The position error estimates for the acoustic data range from 2.6 to 7.9 metres at 90% confidence. Values are assigned to the position accuracy attribute of each sounding in processing software. For the Lidar depths, a value for position accuracy of +/-5 metres and +/-0.2 metres for depth measurement accuracy for all measurements at 1 sigma was supplied by Terra Surveys. With an estimate of errors due to tides and sea-surface roughness of 0.1 metres each, and bringing all estimates to 90% confidence we get a priori position accuracy of +/-11 metres and a priori

Test Area	Est. (90%) depth error	Est. (90%) position error	Number of points
A	0.35	11	1033
В	0.35	11	1062
С	0.35	11	1011
D	0.35	11	931
E	0.35	11	998

reduced depth accuracy of +/-0.35 metres. A summary of the Lidar data sets is given in Table 3.

Table 3 - Summary of Lidar depth data for five test areas.

6. RESULTS

The comparison of the gridded values from the acoustic and Lidar data sets, using the variograms computed from the dense, along-track acoustic profile data was performed by subtracting the Lidar depth estimates from the acoustic estimates at the common grid intersections. 1600 points were available at 20 metre spacing for all areas except area "C" which only had 1480 comparisons because of the lack of acoustic data in the very shallow water close to the shore of Camping Island. Scatter plots of the differences for the test areas are shown as a function of depth in the next five figures. The horizontal and vertical scales for all the scatter-plots are the same, so an indication of the variability in the noise levels and the depth range for each of the test areas can be easily seen. The figures are shown in order of increasing depth so that the change in the depth bias is apparent. Note that there was no test area with depths in the range of 9 to 16 metres, which is unfortunate since this is the depth range of most importance for shipping in Dolphin and Union Strait.

Compare the biases of the test areas where the depths are common - in particular where the minimum depths of area "D", Figure 7, of about 3.7 metres are the same as the maximum depths of area "C", Figure 6 and where the minimum depths of area "E", Figure 10, of about 24.6 metres are the same as the maximum depths of area "B", Figure 9. Notice that the offsets of both of these pairs of test areas is about 0.4 metres. Because of the direction of the flight lines for the Lidar (more or less parallel to the shores of Dolphin and Union Strait - see Figure 2) Areas "B" and "C" were surveyed at more or less the same time - between 19:23 and 21:10 on Julian day 229, 1990. This constant offset for these two areas suggests that a large error exists in the predicted tides for this day. By adjusting these two data sets by the difference between predicted and observed tides (recall the gauge at Bernard Harbour worked until day 230) - 0.14 metres - the fit of the data to a straight line is much improved.

100



FIG. 6.- Depth difference scatter-plot for test area "C".



FIG. 7.- Depth difference scatter-plot for test area "D".

101



FIG. 8.- Depth difference scatter-plot for test area "A".



FIG. 9.- Depth difference scatter-plot for test area "B".



FIG. 10.- Depth difference scatter-plot for test area "E".

7. DISCUSSION

A summary of the statistical parameters from the above comparisons is given in Table 4. The estimated difference error, obtained by taking the mean of the root-sum-square (RSS) of the estimated errors from the Lidar and acoustic data sets and scaling to 90%, is always greater than the 90% deviation calculated from the depth differences themselves. This inequality suggests that the error estimates, which combine error contributions from depth and position measurement, reduction and interpolation (due to ruggedness of the seafloor), are pessimistic. Since the differencing operation correlates both sets of data, there is no way of establishing which of the error components (Lidar depth, position, Acoustic depth, position or any combination) may have been overestimated.

Test Area	average acoustic	average Lidar	depth 'diff.	deviation of diffs.	est. Lidar errors (68%)	est. acoustic errors (68%)	est. diff. error (90%)	correl
Α	18.56	18.9	-0.34	0.36	0.28	0.08	0.47	0.97
В	23.73	24.06	-0.47*	0.21	0.17	0.05	0.29	0.97
С	2.85	2.53	+0.18*	0.24	0.16	0.06	0.28	0.93
D	7.23	7.35	-0.12	0.30	0.26	0.06	0.44	0.98
E	27.95	28.7	-0.72	0.51	0.36	0.16	0.65	0.97

Table 4 - Statistical summary of Acoustic/Lidar depth comparisons @ 90% confidence,

* depth differences adjusted for difference in predicted and observed tides at Bernard Harbour.



FIG. 11b.- Test Area "E" Mean Depth: 28 m.

A comparison of the acoustic traces in Figure 11 a and b to the scatter plots in Figures 9 and 10 clearly shows the relationship between noise in the differences and roughness of the seafloor. Herein lies the reason that position and interpolation errors must be considered when assessing the noise in the differences between two data sets. The acoustic traces are shown at the 0-50 metre scale. Note the differences in the shapes of the distributions (11c, 11d).

Figure 12 shows the differences plotted as a function of depth, together with lines denoting the 90% spread of each set of differences about the mean. If we choose to ignore the suspect area "C" (suspicious perhaps because of errors in predicted tides or due to less sophisticated signal processing) and force the linear regression through the origin, as is typical of other Lidar/acoustic calibrations [e.g. BILLARD, 1986], weighting the remaining areas using the inverse of the variance of each set we get the following equation:



Corrected Lidar depth = present Lidar depth* 0.98

FIG. 12.- Bias (Acoustic depth - Lidar depth) and 90% bounds of difference scatter from five test areas.

The correlation of this regression is quite good: 0.95. The resultant errors introduced to the corrected depths for the average depth of the test areas is as given in Table 5. As can be seen, the errors are very small for three of the test areas chosen for the fit, but the shoal bias error introduced in area "C" is large, even though it is biased on the side of safety. Further, there is a deep bias introduced in area "E".

(1)

Area	Measured Lidar	By Equation 1	Acoustic	Error
C	2.67*	2.62	2.85	-0.23
D	7.35	7.21	7.23	-0.02
A	18.9	18.52	18.56	-0.04
В	24.2*	23.72	23.73	-0.01
E	28.7	28.10	27.95	0.15
RMS				0.133

Table 5 - Errors introduced using linear regression (negative values indicate a shallow bias). * corrected for differences in observed and predicted tides at Bernard Harbour.

The drawback to this approach is that if the areas "B" and "C" are biased as a result of predicted tide errors for day 229, 1990, then all the other depths measured on this day will absorb some of the error. For this reason, the depth differences across all of Dolphin and Union Strait (in areas other than the five test areas) where Lidar and acoustic data overlap, in particular in the identified corridor in Cache Point Channel, will be examined in order to test the validity of this method. Discussion of this comparison will be presented in a sequel to this paper.

SP-44 Specifications and CHS Survey Standing Orders for Depth Accuracy

SP-44 , 3rd Edition, 1987, IHO standards for Hydrographic Surveys, Part 1.C.1 defines the 90% limit on the total error in measuring depths as +/- 0.3 metres for depths in less than 30 metres of water. It also suggests the tidal reduction errors should not exceed this same value. By propagation of errors then, the differences in check soundings should not exceed +/- 0.6 metres in areas of flat or gently sloping bottom. Examination of difference deviations in Table 4 shows that no area has a value which exceeds this precision criterion, ignoring the bias component. According to IHO specifications, the two systems agree to within an acceptable tolerance if the biases can be properly removed. CHS Survey Standing Orders (SSO's) on the other hand state that the tide reduction errors should be half the size of the measurement error, or +/- 0.15 metres, although no confidence level is quoted. This results in checkline agreement requirement of +/- 0.5 metres at 90%, but no regard has been given to the bottom roughness. Area "E" just fails to meet CHS SSO's and Figure 11b shows us why this is so - the seafloor is rugged.

If the estimates of errors for the acoustic measurements are reasonable, then since the estimated errors of the differences (see Table 4) are consistently larger than the 90% deviation calculated from the differences themselves, it would suggest that the precision of depth measurement of Lidar is slightly better than the estimate provided by Terra Surveys. The author believes that in area "B", since the seafloor is quite flat, the position errors contribute only a small portion to the total error in depth. In fact, including a position error component in Hydrostat only caused a 12.5% rise in the estimated errors for the interpolated grid nodes whereas including a depth error component caused a 50% increase in predicted error. The fact that the estimated errors are about 30% larger than the error of the actual differences suggests that the precision estimate may be about 20% too large. So Lidar measurement capability would seem to be at about the +/-0.16 metres level (68%). If a sea-surface component of 0.1 metres is included with the measurement errors and the values are brought to 90% confidence, the Lidar depth measurement precision is at the +/-0.28 metres level, which meets both IHO specifications and SSO's. The problem, as the author sees it, is not that Lidar has the measurement <u>precision</u> to meet IHO specifications, but that because of calibration problems a similar <u>accuracy</u> may be more difficult to meet. The biases in the Lidar calibration are disturbing and need to be properly resolved. If the root-mean-square (RMS) error of the differences between acoustic and calibrated Lidar is calculated (i.e. the standard deviation of the measurements from a zero mean and not from the mean calculated by the measurements - the bias) and the values are brought to 90%, all areas meet IHO specifications, but area "E" still fails to meet SSO's (see Table 6).

Test Area	Residual Bias	90% deviation	90% RMS of diffs Acoustic - corrected Lidar
С	-0.23	+/-0.24	+/-0.44
D	-0.02	+/-0.30	+/-0.30
A	-0.04	+/-0.36	+/-0.36
N	-0.01	+/-0.21	+/-0.22
E	0.15	+/-0.51	+/-0.56

 Table 6 - RMS error of differences between acoustic and calibrated Lidar depths with bias and noise components.

7. CONCLUSIONS AND RECOMMENDATIONS

It is unfortunate that an area was chosen for calibration of Lidar measurements which has a complex and not well understood tidal signal. It is more unfortunate that only one tide gauge was installed for the 1990 Lidar survey and that this gauge failed after only a few days of operation. Perhaps, considering the expense of such an operation, it will be considered cost effective to install several gauges for increased accuracy and reliability of future Lidar surveys in the Arctic (or any surveys for that matter). It would also be worthwhile to eliminate all the other potential sources of systematic biases, such as the temporal aspect of the seafloor and the differences in tides. This can be done by performing both acoustic and Lidar measurements nearly simultaneoulsy. It will be interesting to see if more sophisticated signal processing can cure some of the bias problems seen in the shallow-water areas and perhaps reduce the noise in the deeper water - or extend the reliable depth capability of the system. Information contained in the Lidar waveform about the back-scatter from the water column, can potentially be used to estimate the Lidar propagation biases. If this estimation can be done accurately and reliably then perhaps Lidar surveys can be performed without the need for acoustic survey support. Further investigation is needed.

Acknowledgment

The author gratefully acknowledges the suggestions given by Michel DAGBERT (GSII), Jim VOSBURGH (Terra Surveys Ltd.) Henry WONG (Phd. Candidate in Electrical and Computer Engineering at Uvic), Alex RAYMOND, Ernie SARGENT, Fred STEPHENSON, Tony MORTIMER and Barry LUSK (CHS-Pacific) and Mike CASEY (CHS-HQ).

References

- BILLARD, B., R.H. ABBOTT and M.F. PENNY. 1986. Modeling depth bias in an airborne laser hydrographic system, Applied Optics, Vol. 25, No. 13, pp. 2089-2098.
- GUENTHER, G.C. and R.W.L. THOMAS. 1984. Effects of propagation-induced pulse stretching in airborne laser hydrography, Proceedings SPIE Ocean Optics VII, Monterey, pp. 287-296.
- KIELLAND, P. and M. DAGBERT. 1992. The use of spatial statistics in hydrography, International Hydrographic Review, Monaco, LXIX(1), March, pp. 71-92.
- MYRES, J.A.L. 1990. Assessment of the precision of soundings, Professional Paper Number 25, Hydrographic Department, Ministry of Defence (NAVY), London, U.K., 41 p.
- VELBERG, P.J. 1993. The accuracy of the depth information of the nautical chart, The Hydrographic Journal, No. 68, pp. 29-35.
- WONG, H. and A. ANTONIOU. 1994. One-dimensional signal processing techniques for airborne laser bathymetry, IEEE Transactions on Geoscience in Remote Sensing, in press.