Dear Anonymous Reviewer #1,

Thank you for reviewing the LiDAR and ultrasonic distance sensor parts of this paper. Given the importance of surface roughness measurements to the study, your valuable assessments of these techniques will contribute to a much improved version of the paper.

1. LiDAR Sensor

In our responses to each of these points we aim to demonstrate that the Leica C10 can be used to measure mm-cm scale features at short range, and that our results are above the noise floor of the system.

a. The reviewer is almost correct – the beam divergence of the C10 is 0.24 mrad. However, we think that the reviewer has mistaken the beam diameter of the C10's laser plummet for the diameter of the scanning laser. The manufacturer report the 'dot diameter' of the laser plummet as 2.5 mm @ 1.5 m range, as mentioned by the reviewer, but they report the fixed-width half-height 'spot size' of the scanning laser as within 4.5 mm from 0-50 m range [Leica C10 Datasheet, 2009]. The scanning laser is more powerful than the laser plummet and has a smaller beam divergence.

We use the relationship of Baltsavias 1999 (Section 3.2.; no equation number) to calculate the laser beam diameter as: $d = 2R\tan(div/2)$, where R = range and div = beam divergence = 0.24 mrad. This relationship is valid when div is small, so that $d << R$. The minimum range in our scans was never less than ~ 3 m and our maximum range was typically ~ 10 m. At a range of 5 m, d = 1.2 mm, at a range of 10 m, d = 2.4 mm; much smaller than suspected by the reviewer.

Licthi and Jamtsho 2006 have shown that some overlap between consecutive lidar samples is actually beneficial because the cross-section of the laser beam's power distribution is Gaussian. Therefore the sampling interval can be smaller than the laser spot size to yield a higher Nyquist frequency. Their theoretical calculations identified an optimal angular sampling interval of 0.859*div. In fact, Pesci et al. 2011 have demonstrated through laboratory experimentation that the theoretical limit of Licthi and Jamtsho is pessimistic and consecutive laser spots can overlap by at least 30%.

Using the criteria of Licthi and Jamtsho, our optimal sample spacing at a range of 5 m = 1.03 mm, and at 10 m = 2.06 mm. If we set up one of our scans, e.g. at the SERF ice tank, so that across-track spacing was 2 mm at 10 m range (close to the calculated optimal point spacing for 10 m range), the angular sampling interval would be ~ 0.2 mrad. The across-track point spacing at 5 m range for this scenario would be 0.99 mm, which is still close to the optimal point spacing at 5 m.

Consequently we don't believe that the laser beamwidth of the C10 scanner (specifically) is a limitation when sampling with a small point spacing at the low ranges used in our study.

b. Undeniably, this is a significant limitation of scanning from low incidence angles and one that we aim to test rigorously in the near future.

For this study, the scanner was typically mounted at 2.5-3 m height, either on a platform and tripod in the field. These details have now been added to the manuscript. For an across-track beamwidth of 1.2 mm at 5 m range, the along-track spot diameter is approximately 2 mm, and for an across-track beamwidth of 2.4 mm at 10 m range, along-track spot diameter is approximately 8 mm. This is following a modified version of the equation in Section 3.2.1. of Baltsavias 1999 (no equation number given). Using the criteria of Licthi and Jamtsho, these values for the along-track spot diameter correspond to along-track optimal sample spacing of 1.7 mm at 5 m range, and 6.8 mm at 10 range. Using the less conservative criteria of Pesci et al 2011, these values correspond to 1.4 mm at 5 m range, and 5.6 mm at 10 m range.

As the reviewer states above, the first laser return stronger than the internal thresholding of the sensor is interpreted as the range. The laser pulse becomes spread in time as it backscatters from an inclined surface. However, the threshold of the lidar sensor defines the peak of the return, so that range is detected from the peak rather than the limit of the time-spread return. Consequently, the xy coordinate of a point on the surface is not within the entire footprint of the laser pulse, but within the pulse-peak of the footprint, which is considerably smaller.

This issue is also not only symptomatic of terrestrial lidar, but also of laser profilers that have regularly been used to measure soil/ice surface roughness at these scales [e.g. Drinkwater 1989; Callens et al 2006; both TGRS]. Particularly at higher ranges, along-track spot diameter can be several times larger than across-track spot diameter, so there may sometimes be some correlated sampling in the along-track direction. For most of the data we present, the full-width along-track spot diameter is approximately 50-200% larger than the across-track spot diameter. This may affect the roughness isotropy and is a limitation of the technique. However, samples only become correlated (i.e. cannot be differentiated) if the peaks of the laser footprints cross, rather than the absolute limits of the footprints.

c. The manufacturer modeled surface precision/noise as approximately 2 mm [Leica C10 Datasheet, 2009]. We have re-tested the precision of the scanner over laminate office floor, using a similar setup to our field measurements, and measured the precision as < 1 mm. Further details are provided below.

d. We acknowledge that not all of the roughness elements that are incorporated in a snow/ice surface, e.g. the tiniest grains of ice, can be resolved within the typical diameter of the laser footprint. It is possible that the strength of the reflection (intensity) could be calibrated with samples of ice of known roughness and homogeneous reflective properties to estimate sub-footprint roughness, but that is beyond the scope of this study.

The smallest surface elements cannot be resolved when our samples are spaced at 2 mm. However, we have shown above that the diameter of the footprint was generally below 5 mm along- and across-track in our scans. In almost all cases the point spacing was below the threshold of Ulaby et al. 1982 for sampling roughness (i.e. no more than 0.1x the microwave observable wavelength, so a point spacing of 5.5 mm at the C-band wavelength of 5.5 cm). It is only at the lowest inclination angles and highest ranges that the along-track spacing (only)

exceeds this threshold. We are only really interested in roughness that is around the 5 mm scale for this study.

e. As the reviewer pointed out, the Leica C10 detects the peak of the first return that has a backscatter value above an instrument-defined threshold and computes a range. The method presented herein makes the fundamental assumption that the main source of the backscatter is from the surface of a medium. This assumption is in line with the assumption that can made in a microwave scattering model (i.e., we assume that all scattering takes place at the rough surface).

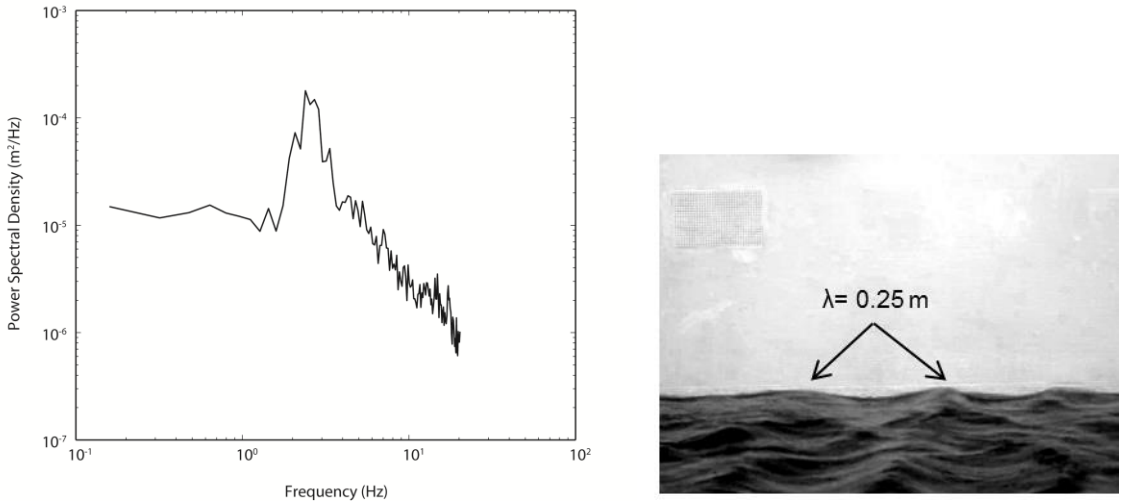
Measurements of the bi-directional reflectance distribution function (BRDF) of sea ice support this assumption. For instance, Perovich (1994) showed that 550 nm light incident on various types of snow and sea ice surfaces from a zenith angle of 60° had a consistent but strongly anisotropic BRDF. There was a significant specular component at 0° azimuth (i.e. in the direction of the illuminating light source; in their case the sun), due to roughness elements of the surface with slopes oriented normal to the incident beam. The reflectance factor then dropped by more than an order of magnitude towards 180° azimuth (i.e. away from the light source). The energy which returns to the Leica C10 by specular reflection from the ice surface would be several times more intense than the energy returning from volume scattering within the upper portion of the ice. This is because the energy returning to the sensor from volume scattering is spread approximately uniformly over the hemisphere. The returning pulse would include a strong defined front from surface reflection and a longer undefined tail from volume scattering.

However, we don't believe that this potential source of error or the small amount of ranging noise introduced, prevent us from using terrestrial lidar for measuring roughness. Many of the existing techniques for measuring roughness have a vertical precision that is considerably poorer than the Leica C10 scanning at low range, including the meshboard, pin profiler, bath chain and laser profiler. We feel that our results provide the best measurements of a range of ice surfaces yet, offering a compromise between precision/sampling resolution and the issues of sampling extent and 1-D profiling that most of the traditional techniques suffer from.

2. Ultrasonic Sensor

We acknowledge the reviewer's concerns regarding the limitations of the TSPC30S1 sensor for measuring wave surface roughness and the inadequate reporting of error as mean error. Motivation for the use of this method lies in difficulties imposed by other non-invasive wave sampling methods when working in a sea ice melt pond environment, e.g. a scanning laser slope gauge (e.g. Bock and Hara, 1995; Plant et al., 1999) requires instrumentation to be installed below the water surface. In terms of selecting the true surface, the TSPC30S1 uses a threshold detection technique, in a timed measurement. A counter starts when the pulse is launched, and stops when the reflected wave reaches the necessary level. Counts are then converted to distance. As suggested the true surface within the estimated 4cm footprint at the half-power point may not be completely represented by the pulse return. This is particularly true for small surface roughness and shorter wavelengths which, within a footprint, result in a messy return and noise concerns due to the sensor's accuracy and repeatability. However returns are most likely coming from the first reflecting point which, from waves of larger roughness and wavelengths >4cm, means that at such a close distance we should contain plenty of the wave signal above noise (one

peak, one trough, etc.). Qualitatively we do see a pk-pk in the sampled data that is consistent with this a returns from wind induced wave trains, though we acknowledge the wave amplitude data in Fig. 9 does not adequately show this. We also see good correspondence between the wave-height spectrum and digital video (DV) samples taken following the grid measurement technique of Scharien and Yackel (2005) – see figures below. By selecting a subsample of these returns from the wave trains and averaging the data to derive our roughness parameters, we are further improving the SNR. For the final version of the paper we will provide an improved analysis of the sensor's capability by: (1) including coincident DV verification data that we collected during the experiment using and (2) accounting for the additional sources of noise described by the reviewer and removing noisy data from the analysis. A more comprehensive description of the methods and discussion of limitations of this technique will be included with similar description/discussion pertaining to the LiDAR data above.



Wave height spectrum (left) for melt ponds derived from 64 wind-wave periodograms across the 2.0 to 11.4 m s-1 U_{10} range, and sample DV frame (right). The peak frequency from the spectrum is 2.4Hz. Using the dispersion relation and deep-water approximation, the phase velocity of the dominant wave is estimated to be 0.65 m s^{-1} and its wavelength 0.27 m. This agrees well with imagery on the right, which shows waves of ~0.25m wavelength.

Bock, E. J., and T. Hara (1995), Optical measurements of capillary-gravity wave spectra using a scanning laser slope gauge, *J. Atmos. Oceanic Technol., 12*, 395–403.

Plant, W. J., W. C. Keller, V. Hesany, T. Hara, E. Bock, and M. A. Donelan (1999), Bound waves and Bragg scattering in a wind-wave tank, *J. Geophys. Res., 104*(C2), 3243–3263, doi:10.1029/1998JC900061.

Scharien, R.K. and Yackel, J.J. (2005), Analysis of surface roughness and morphology of first-year sea ice melt ponds: implications for microwave scattering, *IEEE Transactions in Geoscience and Remote Sensing, 43*(12), doi:10.1109/TGRS.2005.857896.