Project no. 265863

ACCESS

Arctic Climate Change, Economy and Society

Instrument: Collaborative Project
Thematic Priority: Ocean.2010-1 "Quantification of climate change impacts on economic sectors in the Arctic"

D1.29 – Report on altimeter sea-ice thickness errors due to ice type, geometry and snow pack effects

Due date of deliverable: **30/04/2013**

Actual submission date: 12/02/14

Used Person/months: 32

Start date of project: **March 1st, 2011** Duration: **48 months**

Organisation name of lead contractor for this deliverable: **UPMC**

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

1. Introduction 3
2. Sources of error in CryoSat freeboard and sea ice thickness data 4
   - 2.1 Getting elevation from CryoSat-2 data 4
   - 2.2 From elevation to freeboard 5
   - 2.3 Lead detection 5
   - 2.4 Radar penetration into snow 7
   - 2.5 Freeboard uncertainty 8
   - 2.6 Sea ice thickness uncertainty 8
   - 2.7 Discussion 9
3. Comparing CryoSat with large-scale in situ measurements 11
   - 3.1 BGEP moorings 12
   - 3.2 Operation Ice Bridge 14
   - 3.3 Airborne electromagnetics 16
   - 3.4 Summary 19

4 Variation of the draft/freeboard relation with ice topography and its impacts 20
   - 4.1 Introduction 20
   - 4.2 Data acquisition 21
     - 4.2.1 Airborne scanning laser 21
     - 4.2.2 Under-ice swath sonar 21
     - 4.2.3 Co-registering laser and sonar swaths 22
     - 4.2.4 In situ measurements of snow, water and ice 23
   - 4.3 Comparing draft and surface elevation 24
     - 4.3.1 Separating deformed and level ice 24
     - 4.3.2 The surface elevation – draft relation 25
     - 4.3.3 R versus draft 27
     - 4.3.4 FY ice ridge 30
     - 4.3.5 Regionally averaged R 30
   - 4.4 Comparison with Cryosat results 33
   - 4.5 Discussion and conclusions 37

5 Estimating the regional importance of deformed ice 39
   - 5.1 Introduction 39
   - 5.2 Multi-beam sonar data acquisition 40
   - 5.3 Surface roughness classification method 43
     - 5.3.1 Separating level and deformed ice 44
     - 5.3.2 Detailed analysis of variograms for regions 44
     - 5.3.3 Classification by patches 46
   - 5.4 Results 48
   - 5.5 Discussion 52

6 Conclusions 54

References 56

1. Introduction

The thinning and retreat of Arctic sea ice is one of the most serious geophysical consequences of climate change. The disappearance of (reflective) sea ice and its replacement by (absorptive) open water is one of the strongest positive feedback mechanisms in global climate, and the reduction in ice extent is already having serious impacts on Arctic ecology and on northern hemisphere weather systems such as storm track trajectories.
Though satellite-borne instruments have monitored the changes in ice **extent** since 1973, our knowledge of the changing **thickness** of Arctic sea ice is far more limited, since it can only be accurately measured on large scales by sonar from underneath the ice. Sonar measurements have been performed since 1958 by submarines of the US and British navies. The demands of the military have meant that these voyages have been widely spaced in time and have covered sparse and differing regions of the ice cover, however, rather than being driven by scientific needs. Submarines often cannot visit the regions of greatest interest, such as coastal regions and shallow seas.

Much effort has therefore been invested in the use of **satellite altimeters**, either radar (Cryosat-2) or laser (ICESat-2, due to be launched in 2016). These measure the small above-water portion of the ice/snow cover (the surface elevation, or freeboard) and transform it into ice thickness on the basis of a series of estimates and assumptions. Such transformation is fraught with problems which have not been resolved, however. Errors centre on the effects of inhomogeneous ice types/roughness or the presence of open water within the radar footprint. In addition, the radar reflection has been shown to occur within the overlying snowpack, instead of the ice-snow interface, as assumed. Lastly the properties of the snow (thickness, density) are not well constrained at footprint scales, though recent airborne surveys suggest that the existing climatology remains valid.

In addition to these instrument effects, we actually know very little about the real variation of under-water ice (the ice draft) with the overlying topography, since detailed co-incident studies of the air-snow, snow-ice and ice-water are very rare.

We are thus at the risk of moving to space-based sea ice thickness monitoring without really understanding either the physical system or the instruments' response to it. The total volume of sea ice forming and its variability are therefore poorly constrained and there are no benchmarks to test or improve the accuracy of ice thickness in GCMs.

One recent improvement has been the advent of CryoSat data correctly processed for sea ice. Prior to the recent availability of the AWI data used here*, the official ESA product had many problems related to a processor optimised for mountainous terrain, rather than the frozen ocean. In this study we thus use the AWI product to examine: (2) the data itself, and the assumptions therein; (3) comparisons with available large-scale in situ validation measurements; (4) how ice draft actually varies with the overlying ice and snow topography; and (5) the prevalence of 'problem ice types' in the Arctic. Finally Section 6 summarises our findings and indicates ongoing research directions.

* Alfred Wegener Institut: http://www.meereisportal.de/cryosat

2. Sources of error in the Cryosat freeboard and sea ice thickness data

Deriving sea ice thickness from the Cryosat Level 1b data consists of two major steps:

1. Estimation of sea-ice freeboard from CryoSat-2 radar waveforms
2. The conversion of sea-ice freeboard into sea-ice thickness with auxiliary datasets

The first step requires the processing of all available CryoSat-2 Level-1b data over Arctic sea ice, while the second step consists of the interpretation of the retrieved freeboard values. Both steps require assumptions and simplifications, which are based on results of validation activities by ESA and partner organizations. These are described in the following sections. It is a major goal of this processor to estimate not only sea-ice thickness, but also the quantification of uncertainties, which arise from of the quality of the Level-1b product and the necessary assumptions and simplifications. The various quantities discussed in this section are shown graphically in Figure 2.1

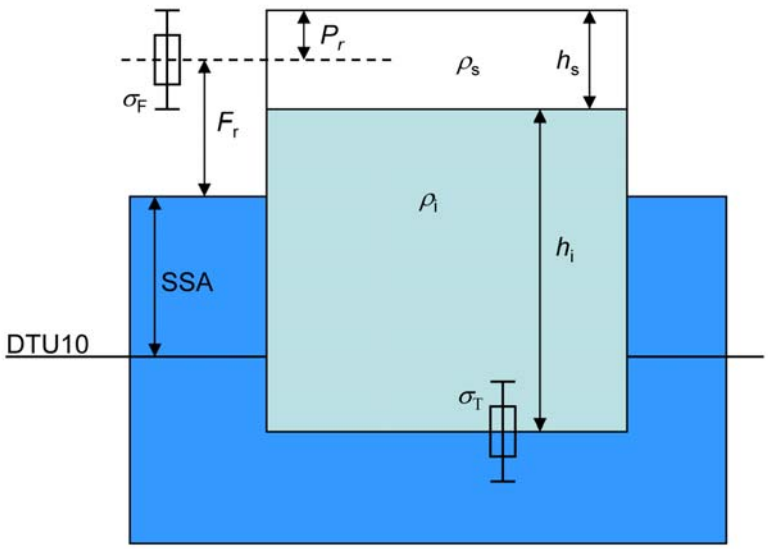

Figure 2.1. Relevant quantities in the process of estimating ice thickness from CryoSat returns. DTU10 is the global mean sea level model used; *SSA* the sea surface height anomaly from this surface, derived by using detected leads as tie points; F_r is the radar freeboard, with uncertainty σ_F; P_r is the radar penetration into the snow layer, of thickness h_s and density ρ_s; h_i is the sea ice thickness, and σ_i its uncertainty; ρ_i is the sea ice density.

2.1. Getting Elevation from CryoSat-2 data

Quality issues with the original ESA CryoSat Level-2 data (Baseline-A) have become clear during the ongoing efforts to define a credible sea ice thickness product. As of 3/2013, the official ESA product is produced using software designed primarily for mountainous terrain. This defines the reflecting surface at the maximum power of the radar return, which is not at all optimised for sea ice studies, and has led to serious problems - dominantly with off-nadir leads being taken for the sea ice surface, producing freeboards of (for example) -23 m.

These shortcomings have led several investigators to implement their own radar waveform tracking algorithms – termed 'retrackers' - which retrieve the distance below the altimeter from the return echo power of the Level-1b (waveform) data files. The AWI CryoSat-2 sea-ice product is chosen here, since it (a) uses an optimised approach for sea ice studies; (b) maintains an overview of the errors and uncertainties involved. The AWI retracker applies a modified TSRA (Threshold-Spline-Retracker-Algorithm, Ferraro and Swift, 1995). This uses cubic spline interpolation and a running mean to interpolate and slightly smooth the waveform. In a second step, the leading edge of the first maxima of the waveform is tracked at a certain threshold level considering the noise floor. The mean distance to the surface is then taken at this retracked position. The TSRA is applied for all surface types (water, sea ice, mixed) to estimate the ellipsoidal elevation *L*.

2.2 From Elevation to Freeboard

In order to obtain sea-ice freeboard, the elevation *L* must be referenced to the local sea-surface height. This is done in two steps:

1. Subtraction of the mean sea-surface height product DTU10 (MSS) from the elevation *L* to remove the major contribution to the sea-surface height changes
2. Automatic detection of open water returns in the ice cover (leads). The elevations of leads define the sea-surface anomaly (SSA), which is the deviation of the actual sea-surface height from the mean sea-surface height.

The first step reduces errors in the detection of the leads and uncertainties in areas where leads are scarce. The second step accounts for deviations (SSA) of the actual local sea-surface height and is calculated for each CryoSat-2 ground track individually.

The subtraction of mean sea-surface height (MSS) and sea-surface anomaly (SSA) from the elevation *L* subsequently yields the height of the sea-ice surface as seen by the radar altimeter. No assumptions or corrections of the radar wave penetration into the snow are made at this point; therefore the result is termed radar freeboard, F_r, and an example plot is shown in **Figure 2.2**, for March 2013.

$$FR = L - (MSS + SSA)$$

2.3. Lead Detection

The accurate retrieval of freeboard crucially depends on the classification of leads in the ice cover. Radar waves over these leads can be automatically classified by the specular shape of the radar waveforms, compared to the more diffuse reflections over sea ice. These specular returns are more "peaky" than their over-ice counterparts and the AWI CryoSat-2 sea-ice processor uses statistical parameters of the waveform stack (namely pulse peakiness *PP*, Kurtosis *K* and standard deviation *STD*) in the level-1b data to automatically detect them. Leads are identified in the AWI sea-ice level-2 processor with the following settings:

$$PP > 40,\ K \geq 40 \text{ and } STD < 4$$

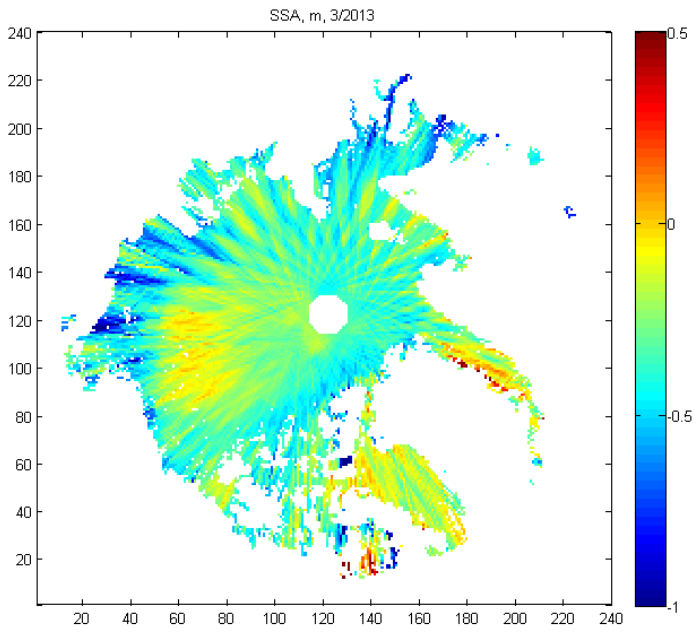

Figure 2.2: Sea surface height anomaly, for March 2013, displaying probable track artefacts in the marginal seas.

However, off-nadir reflections from leads can still dominate the echo and hence bias the range retrieval (known as 'snagging'). Those echoes are detected and subsequently removed by calculating a modified *PP* on both sides of the power maximum of adjacent waveforms, the 'left' (*PP l*) and 'right' (*PP r*). The left peakiness should be high if the lead is in nadir position, though some minor contribution at the sides might be present (*PP l* > 40 and *PP r* > 30). Echoes from ice floes exhibit a lower peakiness at the tail of the waveform and therefore are identified by *PP r* < 15 and *K* < 8. After classification for each CryoSat-2 profile, the elevations of the leads are interpolated and smoothed over a 100 km window.

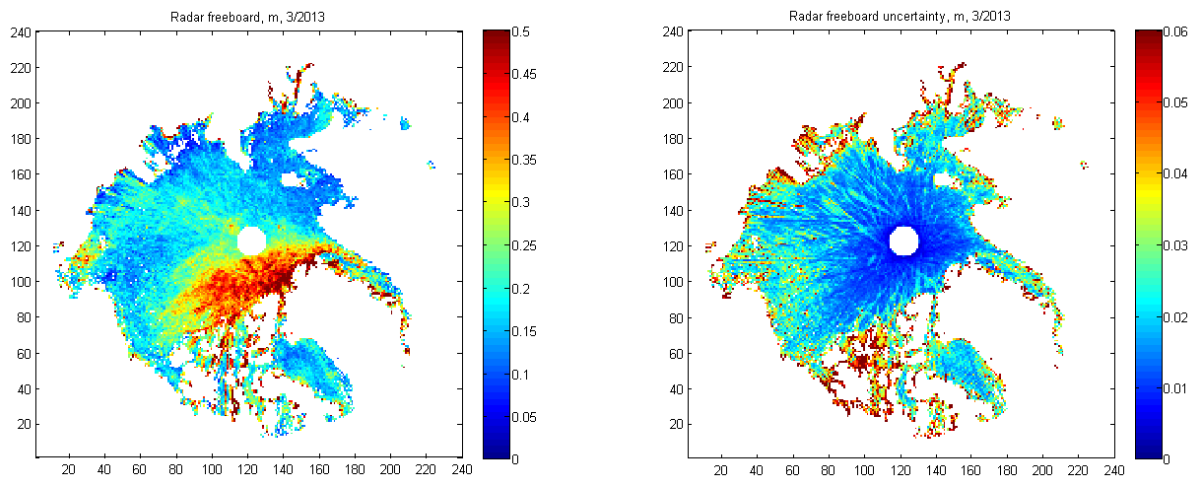

Figure 2.3: Example radar freeboard (left) and its uncertainty (right) from the AWI processor, here shown for March 2013.

2.4. Radar Penetration into Snow

Complex regionally- as well as temporally-varying physical properties (temperature, density, roughness) of the snow layer above Arctic sea ice are a source of uncertainty for the interpretation of the radar freeboard in terms of its location at or above the snow-ice interface. In dry and cold conditions, the main reflector of Ku-Band signal of CryoSat's radar should theoretically be the snow-ice interface. Results from validation experiments *e.g*. by ground radar observations or comparisons between airborne laser and radar altimeter (Hendricks *et al*., 2010, Willatt *et al*., 2010, Willatt *et al*., 2011) have shown that the retracked elevation is often within the snow layer, or sometimes at the air-snow interface (*i.e*. with no penetration of the snow layer at all). Since validation data regarding penetration are sparse at basin scale, we apply a simplified parameterisation which takes into account seasonal changes: an additional penetration factor *P* is introduced, which describes the penetration of the radar reflection horizon. For the entire Arctic we assume a maximum penetration $Pmax$ of 23 cm from November till April, 11 cm for October and May, and 0 cm from June till September. The penetration factor is set to not exceed local snow depth.

Note: Setting an Arctic-wide penetration factor very likely does not reflect the regional different snow conditions and is one important factor of uncertainty for CryoSat-2 Arctic freeboard. However, the knowledge of regional and temporal variation of CryoSat-2 radar penetration into the snow layer is limited by the lack of airborne validation data in early winter and the central and eastern part of the Arctic Ocean. Future revisions of the AWI CryoSat-2 sea-ice level-2 processor shall contain an improved handling of this parameter in freeboard retrieval.

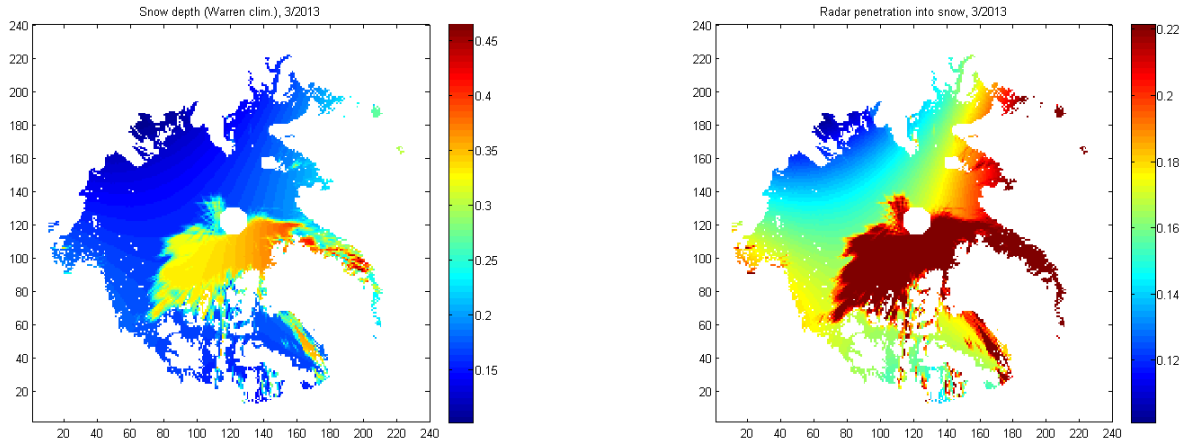

Figure 2.4: Snow depth (LHS, from the Warren climatology) and radar penetration depth (RHS), for March 2013, when the snow can be relied upon to be cold and radar penetration is at its maximum. For the FY ice area (snow depth <20 cm), the radar is taken as penetrating to the ice/snow interface. Thicker snow in the MY ice region north of Greenland displays the 'fixed' penetration depth of 23 cm.

2.5. Freeboard Uncertainty

In summary, the uncertainty budget of the freeboard retrieval depends on:

1. The quality of the Level-1b radar echoes
2. The skill of the SSA detection algorithm, that references the elevations to the actual sea-surface height
3. The knowledge of radar penetration into the snow

Assuming uncorrelated errors and thus standard error propagation, the uncertainty of each individual freeboard measurement from CryoSat-2 can be described as:

$$\sigma_F{}^2 = \sigma_L{}^2 + \sigma_{SSA}{}^2 + \sigma_P{}^2$$

with σ_L the uncertainty for the range measurements (radar speckle), σ_{SSA} the uncertainty of sea-surface anomaly and σ_P the uncertainty of the approximation of radar penetration. This uncertainty is likewise calculated for each data point individually. Except the SSA uncertainty, which is a function of the standard deviation of lead elevation and the distance to the next lead, all other uncertainty contributions are set to a fixed values.

2.6. Sea-Ice Thickness Uncertainty

The estimation of sea-ice thickness follows the same principles as the estimation of freeboard uncertainty for each data point. A more detailed description and weighting of the different sources of uncertainty is given in Giles et al. (2007).The thickness uncertainty depends on the uncertainty of freeboard, snow and ice density as well as snow depth. The seawater density does not vary significantly and is neglected in the uncertainty budget.

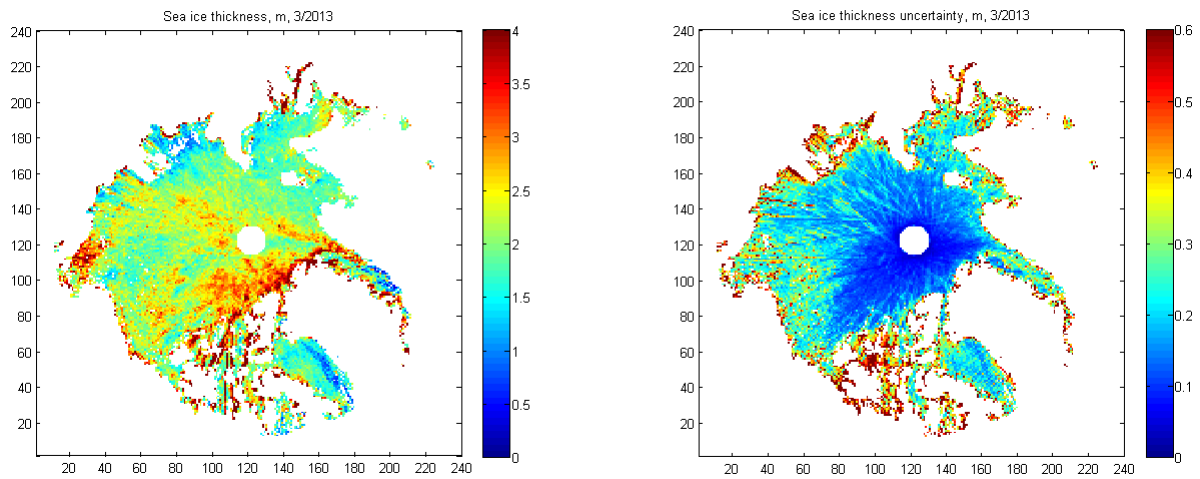

Figure 2.5: Sea ice thickness (left) and uncertainty (right), for March 2013.
Ice type and concentration

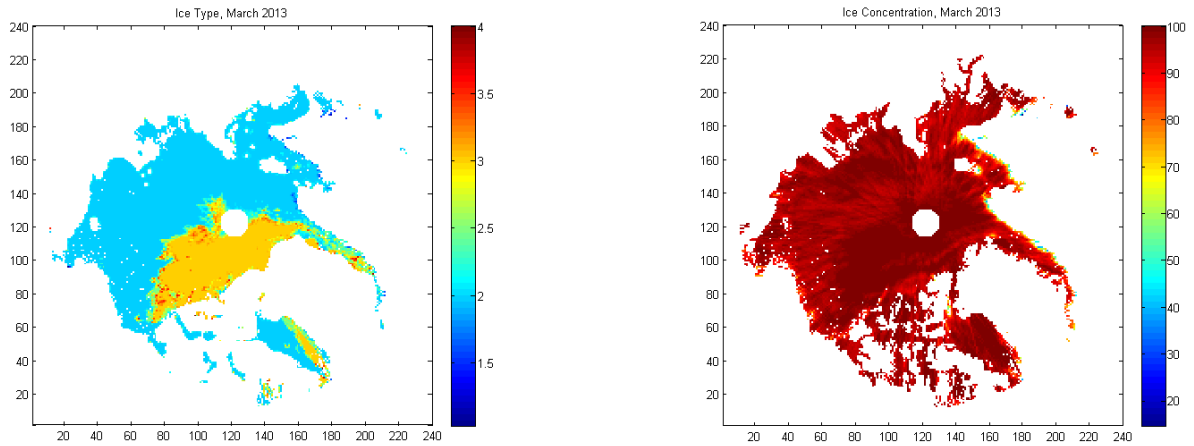

Figure 2.6: Ice type (left) and ice concentration (right) for March 2013. Ice concentration field shows some track-based artefacts.

2.7. Discussion

The uncertainty fields from the AWI product appear to give some confidence in the ice thickness result, since the ice thickness uncertainty is generally less than 30 cm over the Central Arctic Basin, only degrading noticeably close to land or in the marginal seas. Similarly, the radar freeboard uncertainty is apparently only around 1-2 cm in the Central Arctic.

These tight error bounds appear unlikely to be true in practice, given the largely unconstrained nature of the applied snow depth (the Warren climatology) and the unknown variation of the radar penetration depth. The error budget presented also ignores issues of the representativeness of the radar freeboard return over deformed ice, discussed in **Section 4**.

Figure 2.7 shows the available timeseries for a point in the Beaufort Sea (73N, 145W), corresponding to the location of the APLIS ice camps, discussed in Section 4. Uncertainty values in ice thickness and radar freeboard are shown as error-bars on those plots. The result for March 2012 shows the effect of a significant error in radar freeboard (red line) when converted to ice thickness (blue line with error-bars). The radar freeboard is generally taken to follow the ice surface (black line), since snow thickness is always below the supposed radar penetration depth of 23 cm. The radar reflection horizon deviates from the ice surface only in late Spring 2011 and 2012 (May), when the snow is taken to be wet.

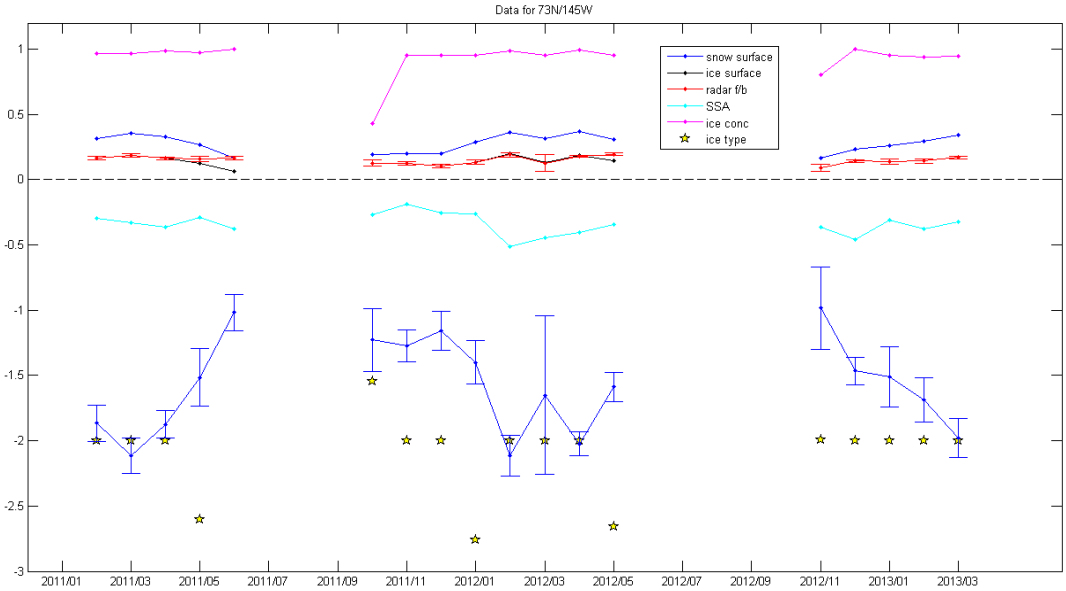

Figure 2.7: Cryosat time series at 73N/145W.

A similar plot from the northern limit of the CryoSat sensor (**Figure 2.8**) shows rather different character. Sea ice thickness uncertainty now looks very low, as a result of the apparently well-constrained radar freeboard (red). The thick snow (max 34 cm) on the MY ice results in the radar reflection (red line) coming from inside the snowpack, not co-inciding with the ice-snow interface (black line) at any point.

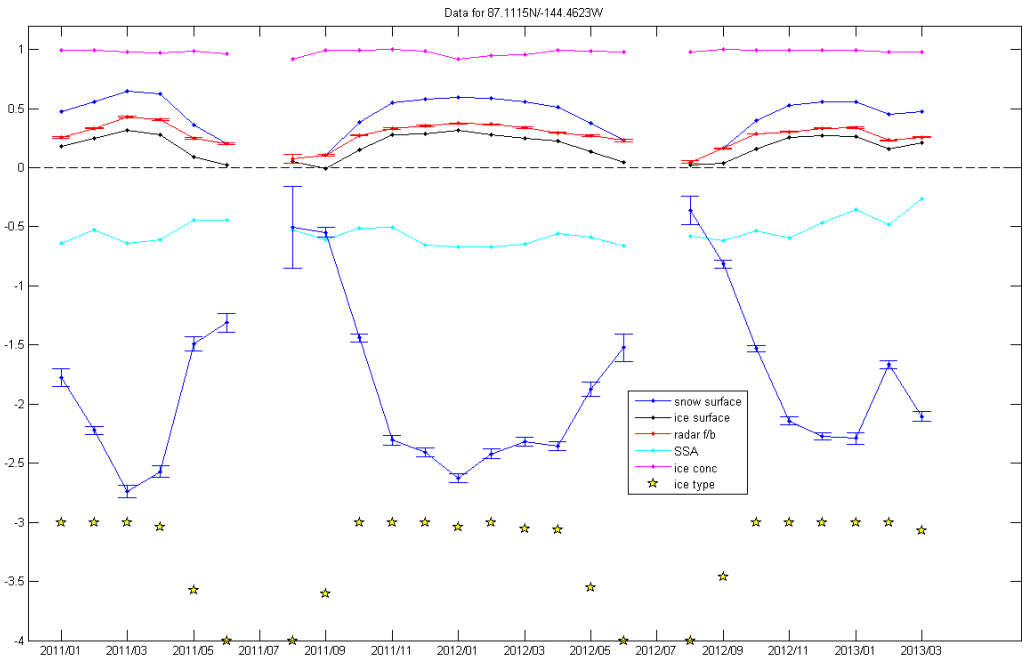

Figure 2.8: Cryosat data for all available months close to the North Pole (87N, 145W).

In the next section, we examine how realistic these stated uncertainties are by comparing the CryoSat results with in situ measurements, co-incident in space and time.

3. Comparing Cryosat results with large-scale *in situ* measurements

Cryosat results are compared with the available large-scale measurements, shown on the map below. These comprise (1) Long-term upward-looking sonar mooring data from the Beaufort Sea Gyre Programme (BGEP), at three locations in the Beaufort Sea, marked by the yellow star symbols; (2) Airborne electromagnetic induction flights, as part of the SIZONet, BGEP and PAM-ARCMIP projects. Tracks are marked as blue lines on the map; (3) Operation Ice Bridge laser altimeter overflights. IceBridge data are grouped into 50km clusters, with central locations marked as red dots on the map.

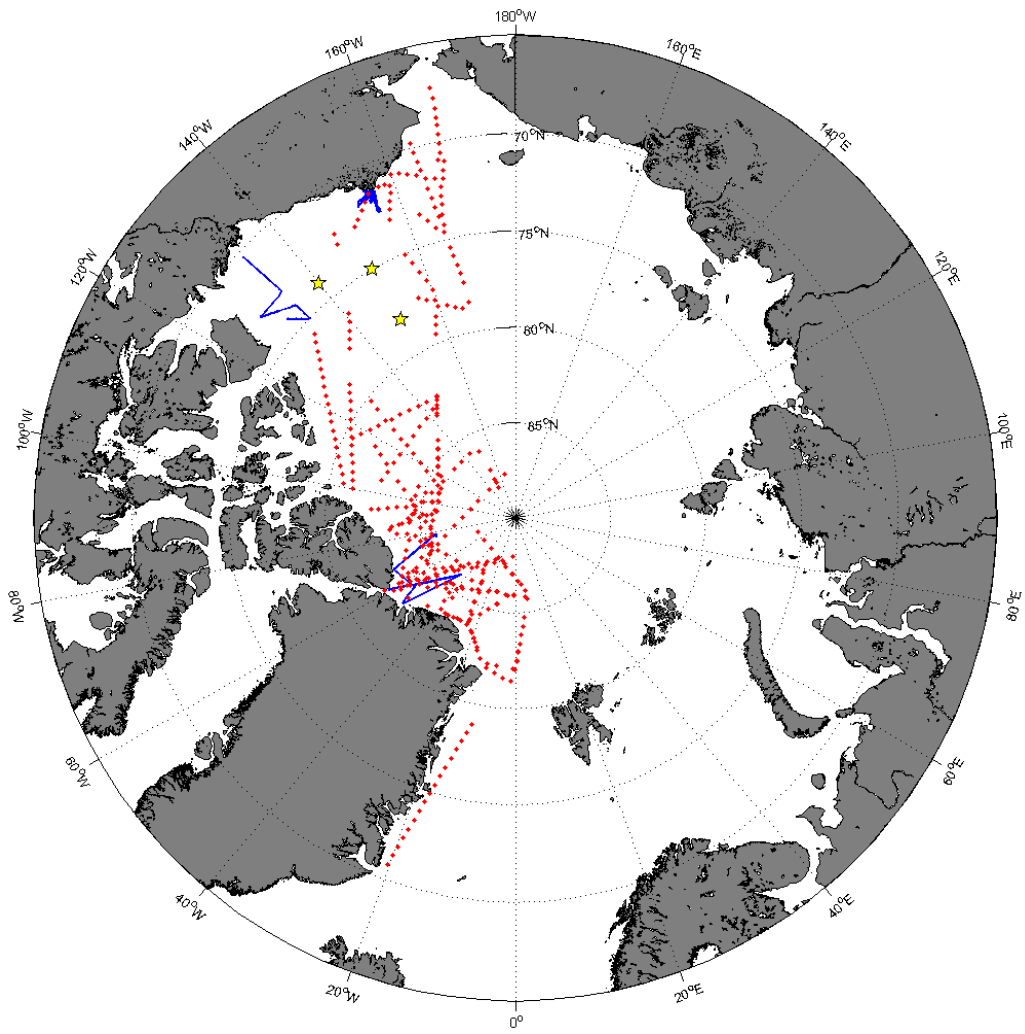

Figure 3.1: Locations of *in situ* ice thickness data for comparison with CryoSat returns

3.1. BGEP moorings

The Beaufort Gyre Exploration Project (BGEP) maintains series of moorings, which include an ASL 420 kHz, 1.8^{o} beam width, upward-looking ice profiler. Only three (A, B and D) of the four moorings operated during the CryoSat data period, and these results are presented here. Data are processed as monthly distributions, derived from more than a million point samples each month, and made available on the Applied Physics Laboratory "Unified Sea Ice Thickness" website http://psc.apl.washington.edu/sea_ice_cdr/data_tables.html

These monthly distributions should be directly comparable with the CryoSat results, extracted at the same locations, and we present these in **Figure 3.2** (a) to (c). During co-incident periods, CryoSat ice draft (plotted in blue, with error bars indicating the ice thickness uncertainty) consistently exceeds the modal or mean values of the monthly distribution seen by the BGEP sonar (black dashed line). The difference is typically 0.5 – 1.0 m, far exceeding the stated ice thickness uncertainty of the AWI product and contrasting with the published draft differences of -0.082+/-0.237m stated in Laxon *et al*. (2013), though they averaged the CryoSat results over a 200km radius around each mooring. Though **Figure 3.2** suggests that CryoSat tracks the mean of BGEP values better than its mode, we might expect the satellite to actually be sensitive to the modal thickness – i.e. the dominant level ice thermodynamic thickness – since this represents the dominant ice type in the footprint, thus dominating the power of the radar waveform return.

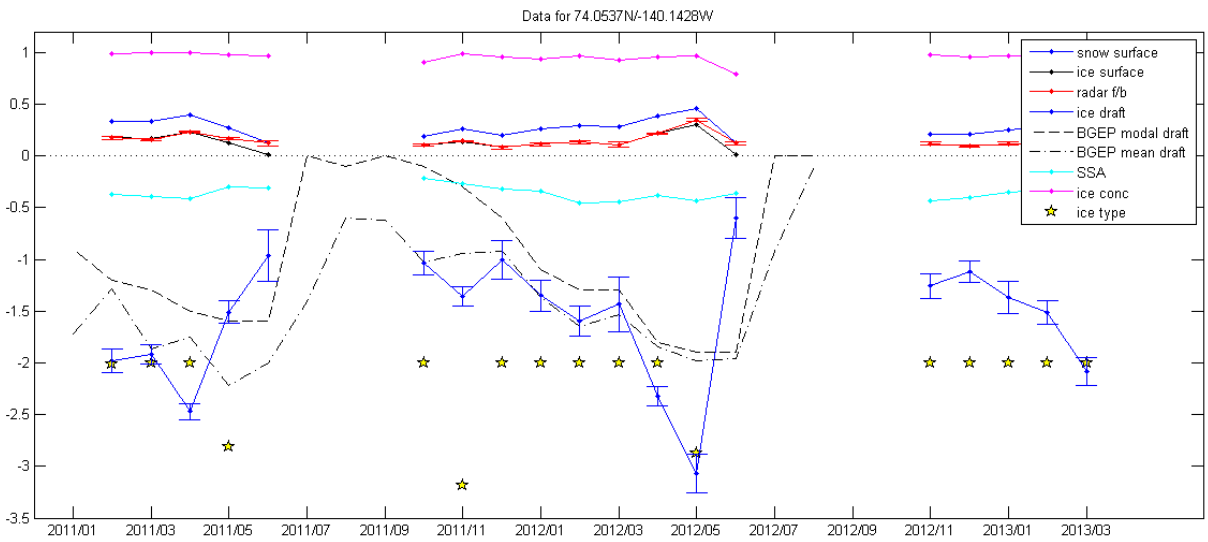

Figure 3.2a: Cryosat results at 74N, 140W compared with the mode (dashed) and mean (dash-dot) of the BGEP 'D' mooring monthly mean distributions.

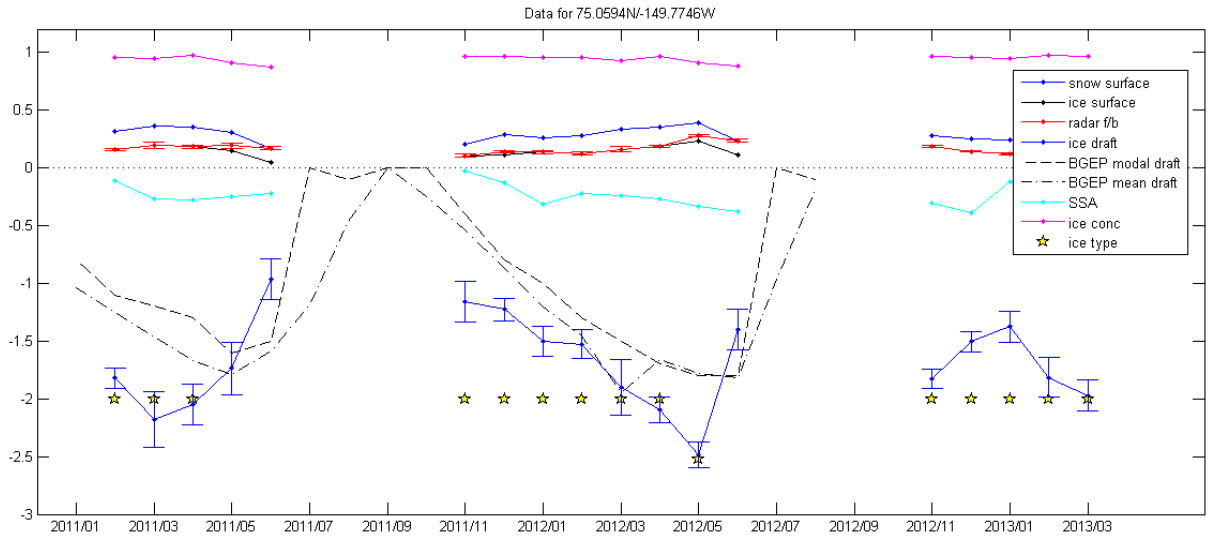

Figure 3.2b: Cryosat results at 75N, 150W compared with the mode (dashed) and mean (dash-dot) of the BGEP 'A' mooring monthly mean distributions.

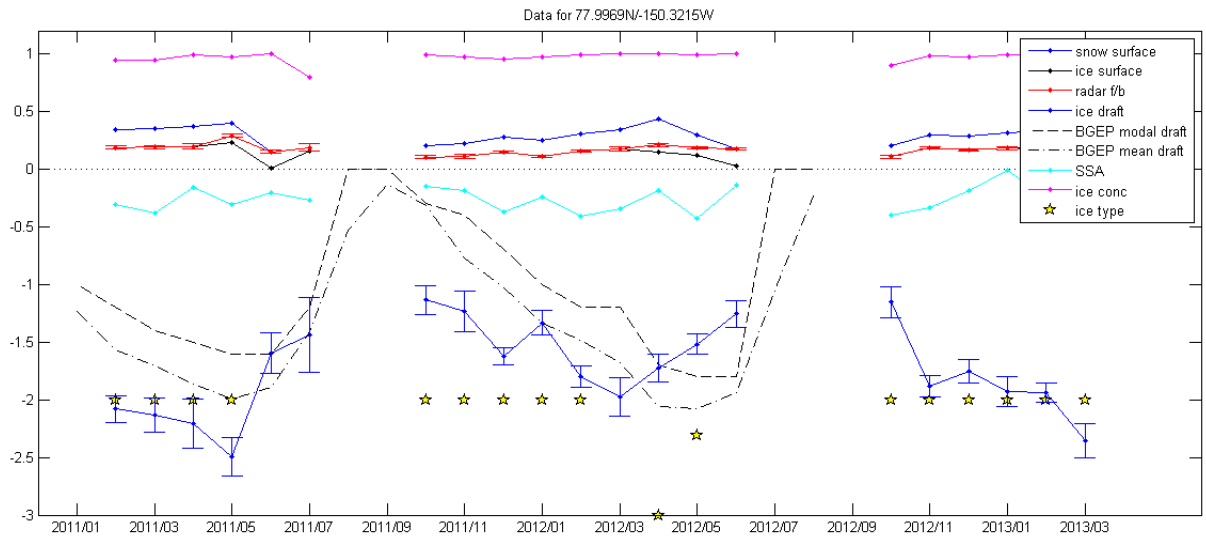

Figure 3.2c: Cryosat results at 78N, 150W compared with the mode (dashed) and mean (dash-dot) of the BGEP 'B' mooring monthly mean distributions.

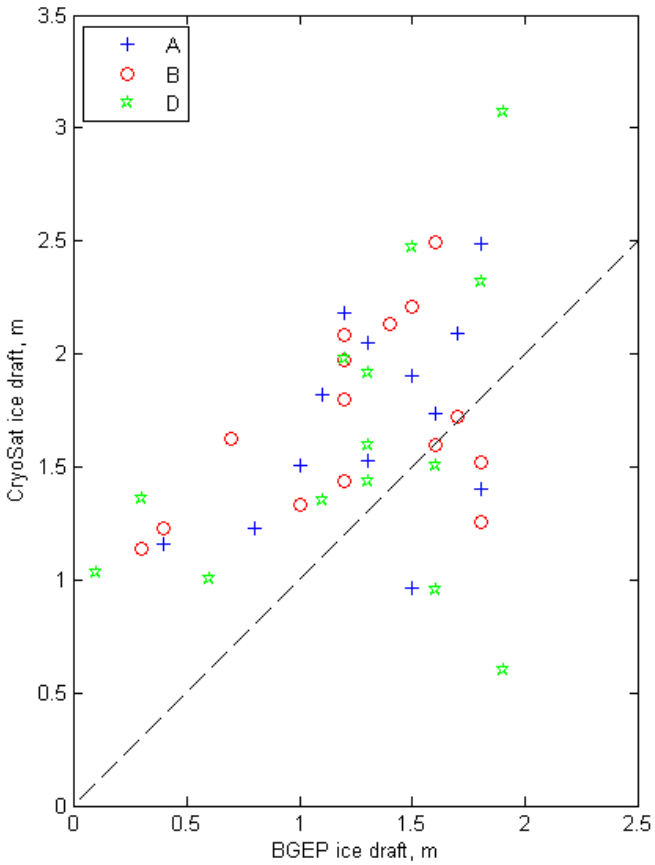

Figure 3.3: Scatter plot of BGEP monthly ice draft modes versus the derived CryoSat ice drafts, for each of the three moorings examined. The dashed line indicates equivalence between the two products.

The consistent overestimation of ice draft displayed by the CryoSat altimeter suggests that the ice thickness uncertainty stated in the AWI product (error bars on Figure 3.2 plots) is significantly underestimating the errors.

3.2. Operation Ice Bridge

Ice Bridge is a NASA-funded programme to monitor Arctic ice thickness from airborne platforms, maintaining data collection during the hiatus between the ICESat-1 and ICESat-2 satellite platforms. The aircraft combines a laser altimeter (giving the snow-air interface) and a wideband (2-8 GHz) snow radar to measure the snow thickness on the ice. This overcomes one of the main limitations of the CryoSat measurements – the unknown snow depth – and should make the IceBridge data more representative of the true ice freeboard.

The 'Freeboard, Snow Depth, and Ice Thickness' data product provided to NSIDC was used to form 50-km clusters, combining data from more than one flight if the flights were less than 10 days apart. The spacing of the point thickness estimates is approximately 25 m. The original data set includes a variable for the uncertainty in the estimated ice thickness that is used to select points with an uncertainty of less than 1 m for very thin ice up to 2 m for ice greater than 4 m thick. The maximum

uncertainty in the point measurements included in the clusters is 2 m. Clusters were required to have 500 or more point samples to be retained and some clusters have as many as 7000 points. The average is 1670 points.

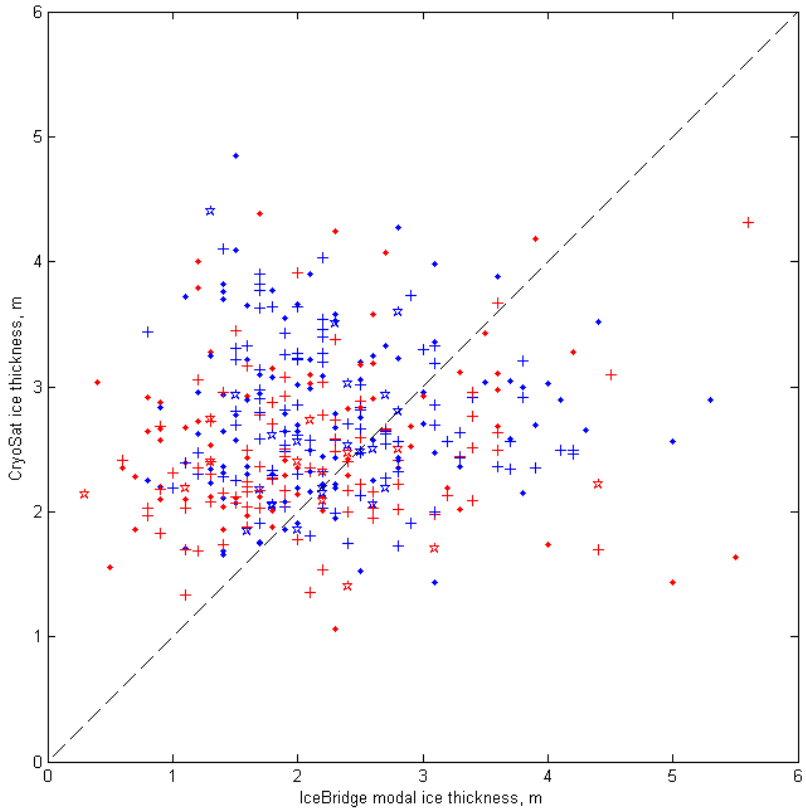

Figure 3.4: Cryosat ice thickness plotted against the modal ice thickness value from IceBridge overflights in March/April 2011 and 2012 (388 distributions). Points are plotted as blue=>above 83N, red=> below 83N. Dots represent distributions with <1000pts, crosses have between 1000 and 2000 points, and stars have more than 2000 points. Little systematic variation is evident.

Data were grouped into 388x 50km clusters and the ice thickness distribution calculated. These modal values are plotted against the CryoSat-derived ice thickness for the same location and month. The lack of correlation is striking, though there is a clear tendancy for CryoSat to over-estimate thickness compared to Ice Bridge (more points to the left of the fit line). No latitudinal or data-sparseness patterns are evident, with points being equally scattered above and below 83N (blue or red symbols respectively), and equally scattered whether the Ice Bridge clusters contain <1000, <2000 or >2000 individual points (dot, cross or star symbols, respectively). Other investigators (Laxon *et al*., 2013) have also found a lower correlation with IceBridge than EM or mooring data.

Figure 3.5 shows the thickness difference per cluster, with the mean (0.44 m +/- 0.17m thicker for CryoSat) marked as a blue line. The RMS difference (i.e. not taking account of the sign) is 0.98m.

Examining the Ice Bridge thickness distributions and the CryoSat thickness value for each individual cluster gives little additional insight into the mechanisms at play, other than the tendency to over-estimate ice thickness in the presence of deformed ice, noted above.

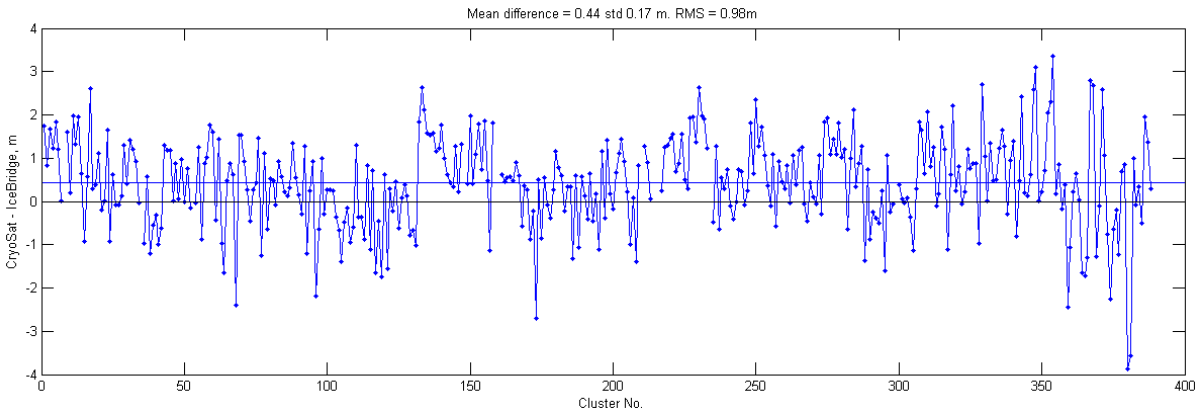

Figure 3.5: Difference between CryoSat and Ice Bridge ice thickness (+ve means CryoSat is thicker).

3.3. Airborne electromagnetics

Data from four campaigns in 2012 and 2013 are presented here: Local surveys north of Barrow, Alaska were carried out in March and April 2012 and 2013, as part of the SIZONet project; a more extended survey was performed north of the MacKenzie delta in April 2013 under the BREA programme; and long transects were performed out of the Canadian base of Alert, into the Lincoln Sea, as part of the PAM-ARCMIP project. The EM survey method has been shown to be very accurate for determining level ice thickness, though its response to deformed, thicker, ice is less well-constrained. Each survey was processed to over 110,000 point values, which were grouped into approximately 50km segments. The modal thickness value was compared to CryoSat results at the same location and time.

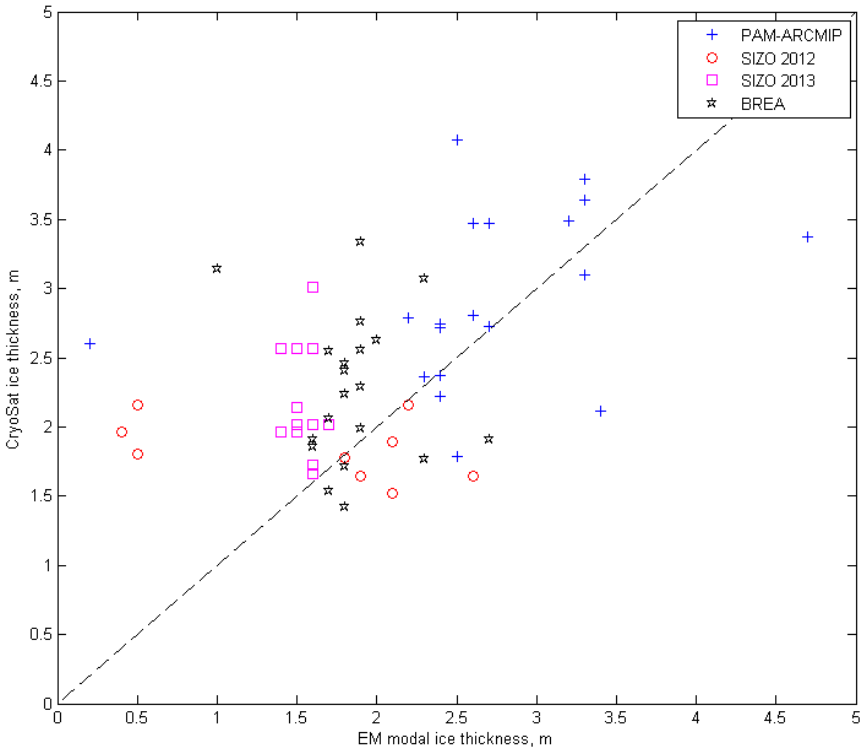

Figure 3.6: Ice thickness values from EM (mode of 50 km distributions) and CryoSat ice thickness compared. Points are coded by campaign.

Looking at the four EM campaigns in detail:

CryoSat generally under-estimates ice thickness for the **SIZONet 2012** survey (red circles), except for three segments where the EM is picking a mode of very thin ice. Examining these PDFs in more detail (**Figure 3.7**) the CryoSat result (colour-coded squares for each case) is struggling in the presence of deformed ice (as seen by the thicker ice modes – blue and red curves), while the green curve suggests a more reasonable pick.

The opposite tendancy is noted for the **SIZONet 2013** survey (magenta squares), which shows CryoSat consistently over-estimating the ice thickness – between 1.5 and 3m – while the modal value according to the EM is fairly uniform at 1.5 m. Examining the distribution shows no reason why this would be the case, since a clear FY ice peak is visible in all the segments. It must be borne in mind, however, that the survey was carried out on 14th April, and April's data are not currently available from the AWI server. It is possible that conditions changed radically between March and April, for instance if MY ice was advected into the area, or the ice was deformed by onshore pressure.

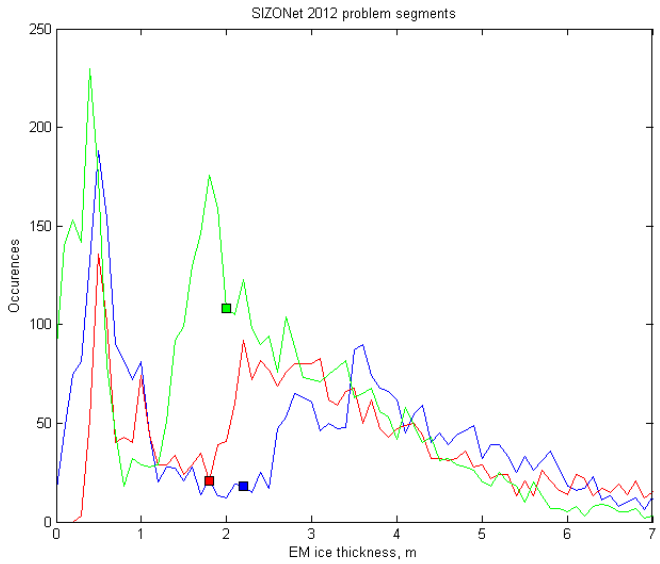

Figure 3.7: Distributions from the SIZONet 2012 EM campaign (curves) with the CryoSat monthly mean ice thickness marked as colour-coded square symbols.

The **PAM-ARCMIP** survey (blue crosses) was mostly over the deformed ice region of the Lincoln Sea, where we might expect CryoSat to struggle. Results show no clear pattern, though there is a tendency for CryoSat to over-estimate ice thickness with respect to the EM results, but, as stated EM is not reliable in this ice type either. It should also be borne in mind that the PAM-ARCMIP survey was carried out between 30th March and 4th April, so may have experienced changed conditions from the March average represented by the CryoSat data. The outlier with very low EM thickness is due to significant open water region (red curve in **Figure 3.8**). In fact CryoSat has picked the 2.6m peak rather well in this case.

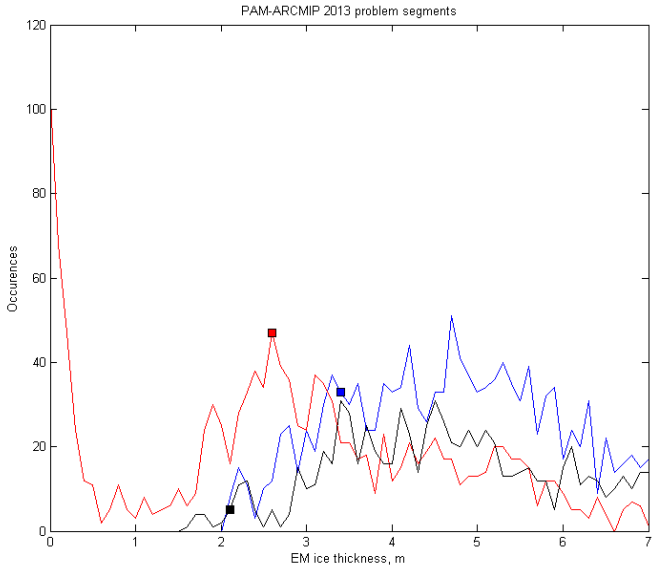

Figure 3.8: Showing the very deformed ice present in the PAM-ARCMIP EM thickness distributions.

Comparison with the **BREA** survey (black stars in **Figure 3.6**) also suffers from a slight time misalignment, having been carried out on 14th and 15th April, while only March CryoSat data are available. This may account for the dramatic over-estimation of the CryoSat ice thickness observed, where no clear reason appears evident from examination of the detailed EM distributions (**Figure 3.9**), which display a bi-modal character with thin ice 0.5-0.8m and thicker ice (2m) modes.

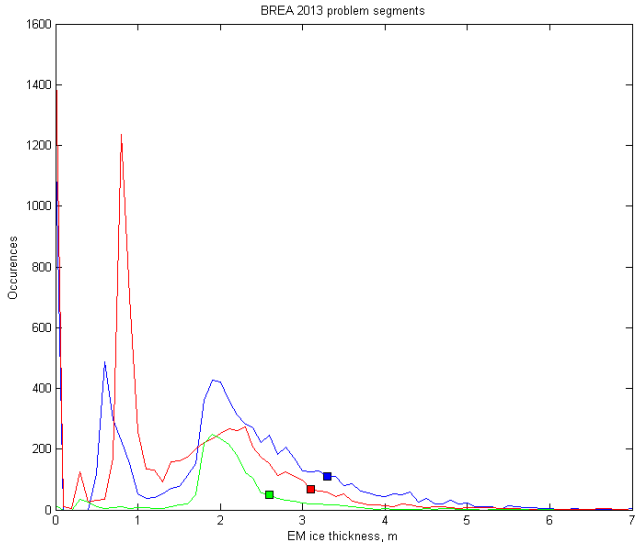

Figure 3.9: Typical EM thickness distributions for the BREA survey.

3.4. Summary

Comparison with the available large-scale survey data clearly suggests that the stated ice thickness uncertainties are exceeded in most valid examples. Only the BGEP mooring data really provides a reliable comparison, in which CryoSat ice thicknesses are generally between 0.5 and 1.0m too thick. The radar freeboard – assumed to be from the ice surface in the presence of a relatively thin snow layer – is perhaps instead coming from within a thicker snowpack.

Ice Bridge data are too scattered to draw any robust conclusions, though a general over-estimation of ice thickness is evident in these data too. Unfortunately these altimeter-based measurements must also transform from a small freeboard to ice thickness, making similar assumptions about material densities and similarly struggling in the presence of deformed ice. The better-constrained snow depth, thanks to the snow radar, although it should improve the Ice Bridge estimates, is difficult to untangle with these datasets.

The available EM data are not ideal as a comparison with CryoSat results, due in two cases to the unavailability of the correct month's CryoSat data, and in another case due to the accepted shortcomings of the EM result itself over heavily deformed ice in the Lincoln Sea. The general underestimation by CryoSat during the SIZONet 2012 campaign shows the instrument struggling in the deformed ice north of Barrow as well.

Broadly, we can conclude that **CryoSat is overestimating ice thickness for undeformed FY ice** (as seen in the BGEP data) and **underestimating deformed ice thickness**. In the next sections we examine the possible reasons behind these errors.

4. Variation of the freeboard-draft relation with ice topography and its impacts

4.1. Introduction

As set out in **Section 1**, CryoSat ice thickness is derived using an often-violated assumption of mono-typic level ice within the footprint, though this is not stated in the literature. Clearly, real sea ice has a mixture of ice types and significant amounts of deformed ice, which may not conform to the simple fixed-density isostatic relations imposed by the processing. We thus examine variations in the relation between freeboard/surface elevation and ice thickness/draft in this section, in the context of a detailed survey over varied ice types, carried out in the Springtime Beaufort Sea.

The motivation for this study was primarily to examine the relation between draft *d* and surface elevation *f* over different ice types. Not enough is known about the detailed relation between these parameters at a time when the ratio, hereafter referred to as $R = d/f$, is becoming a critical parameter as we move towards spaceborne methods of determining ice thickness, such as ICESat and Cryosat-2. These altimetry techniques measure a small number (surface elevation) and multiply it by the poorly constrained R+1 to determine ice thickness. Clearly, such a technique implies significant errors if R is not well known, and R is far from being a single value. Most significant is its variation with snow thickness. Unfortunately, the only time this is less problematic – the summer, when all the snow masking the ice surface has melted – is also the time when radar altimeters cannot determine ice freeboard due to the presence of liquid water on the ice. The chosen ice and water densities also play a role, as does the prevalence of deformed ice, a mixture of ice types within the altimeter footprint and other important, radiometric, issues highlighted in recent work [*Tonboe et al.*, 2009; 2010].

The relation between top- and bottom surfaces is best understood by carrying out detailed co-incident surveys of both surfaces, and this was done as part of the 2007 Applied Physics Laboratory Ice Station (APLIS), using AUV-mounted upward-looking sonar from below and scanning airborne laser from above. The camp was staged from 1st -15th April in the Beaufort Sea north of Prudhoe Bay, Alaska, at approximately 73°N, 145°W, in an area of dominantly 100% ice concentration - as seen by passive microwave satellites - though there were frequent small leads in the immediate area. The AUV measurements presented here were carried out on the 12th and 13th April, with a single laser overflight occurring on the 13th April. AUV operations at the site are detailed in previous publications [*Doble et al.*, 2009; *Wadhams and Doble*, 2008]

Early studies noted that the relation between surface elevation and thickness was best fitted using a thickness-dependent R [*Ackley et al.*, 1976]. Wadhams *et al*. [1992] determined R for multi-year ice in the central Arctic as 7.9 ±0.6. They noted that the laser elevations under-estimated the proportion of ice away from the mode (younger ice and ridged ice) while exaggerating the proportion close to the mode (3-6 m). The pattern was qualitatively explained with reference to the snow cover: thick ice can carry a relatively low snow load compared with thin ice – therefore very thin ice is transformed too thick, while very thick ice (thicker than the modal thickness) is transformed too thin – peaking the distribution. An elevation-dependent R value was therefore suggested. Spatial variability due to ice thickness was expected, with seasonal variability due to snow cover accumulation and melt. *Forsstrom et al*. [2011] found a value of 5.0 ±1.5 in the Fram Strait region in autumn. Investigations in

the Lincoln Sea (north of Greenland) determined *R* as 5.3 for FY ice (implying 16 cm snow cover) and 4.7 for MY ice (44 cm snow cover) [*Haas et al.*, 2006]. They noted that any dependence of *R* with surface elevation needs to be tuned for ice type.

4.2. Data acquisition

The survey was centred on a first-year pressure ridge, which had been observed to form on the 2nd April 2007, shortly before the APLIS camp was established, and was thus poorly consolidated. The ridge divided a region of first year (FY) ice from a refrozen lead. Multi-year (MY) ice floes and FY rubble fields were also present in the area. The camp was located on one MY floe, 1 km distant from the study ridge.

4.2.1. Airborne scanning laser

The ice-plus-snow surface elevation was surveyed on April 13th (co-incident with the AUV sonar acquisition) using a Riegl LMS-Q140i-60 scanning laser profilometer mounted in a Twin Otter aircraft operated by Ken Borek Air Ltd. Data were collected as part of the larger National Space Institute (DTU Space) Spring 2007 campaign. The laser scanner measures with a horizontal resolution of 1×1 m at a flight height of 300 m and a ground speed of 250 kph. The across-track swath width is roughly equal to the flight height. The vertical accuracy is in the order of 10 – 20 cm depending primarily on errors in the kinematic GPS-solutions, due to long baselines [*Krabill et al.*, 1995]. More information on the airborne instrumentation and the system setup can be found in *Hvidegaard et al*. [2006].

If the height of the sea surface is known, the surface elevation can be found directly from the measurements. Here, a geoid model is used as a first approximation of the sea surface height. However, due to tidal errors, ocean dynamic topography and measurement errors, it is necessary to implement a “lowest-level” filtering algorithm. The algorithm selects the lowest heights along the geoid-reduced laser measurements. These heights are assumed to be open water leads, or leads covered with thin ice, and are thus estimates of the instantaneous sea surface height. The distance between such points is typically 5 km, though in the area of the APLIS camp, numerous cracks or small leads were present. A smooth curve is fitted through the “lowest-levels” points by using a least-square collocation (optimal estimation) function with correlation length 10 km and root-mean-square noise of 0.1 m. The resulting curve is an estimate of the instantaneous sea surface height, and the ice+snow surface elevation can be found accordingly [*Hvidegaard et al.*, 2006]. First results, using a similar method, are described in *Hvidegaard and Forsberg* [2002], where the absolute accuracy of the surface elevation is estimated to be in the order of 10 - 15 cm. The geoid model used in the processing is an updated Arctic geoid model, derived by spherical FFT methods from the Arctic Gravity Project terrestrial data [*Forsberg and Kenyon*, 2004] and GRACE satellite data. More details can be found in *Forsberg and Skourup* [2005].

The laser-derived surface elevation data were used as the reference frame for other measurements, since they were essentially synoptic over the widest area and geo-referenced. Maximum observed surface elevation over the co-incident area was 3.2 m and the median surface elevation was 46 cm. The probability distribution function was unimodal, with a 40 cm mode.

4.2.2. Under-ice swath sonar

Ice drafts were obtained using a Geoacoustics Geoswath 500 kHz inferometric sonar, mounted on a Gavia AUV, owned and operated by the manufacturers, then Hafmynd Ehf, now Teledyne Gavia Ehf. The vehicle was run inverted (*i.e*. rolled through 180°) and the sensors, normally configured to look

downwards at the seafloor, thus looked upwards at the undersurface of the ice. A detailed description of the vehicle and sensors is given in *Doble et al*. [2009].

The vehicle was run at depths between 20 and 30 metres, to be sure of clearing the deepest ice features in the survey area, giving a usable sonar swath width of around 80 m (*i.e*. 40 m either side of the vehicle track). The vehicle was run tethered for safe and simple retrieval from the deployment hole, and thus only short missions were feasible. In all, 21 runs were performed, each consisting of an outward track up to 400 m from the hole, a wide turn and return to the hole. Missions were run in a star pattern, with the aim of achieving full coverage of the area out to a given radius, over the course of two days.

Data were processed and binned to 0.5×0.5 m, calculating a weighted mean of the values within each bin using the *Geoswath Plus* software supplied by Geoacoustics. Bins typically contained 20 individual range-angle solutions, each with an estimated draft error of ±15 cm. Error for binned values reduces by the square root of the number of contributions, giving ± 3 cm for draft error in a bin. **Figure 4.1** shows the resulting mosaic of drafts rendered in three dimensions, with various ice types marked and with the probability density function (PDF) inset. Modal ice draft was 1.65 m, reflecting the dominance of level FY ice in the dataset, with a secondary mode at 40 cm, due to refrozen leads. Maximum draft, attained in the FY ridge, was 16.8 m.

4.2.3. Co-registering laser and sonar swaths

The laser overflight was used as the 'master map' to which each AUV sonar swath was matched. Features were identified on the laser surface elevation and on each sonar swath which could reliably be assumed to be co-located on their top and bottom surface, dominantly the edges of floes and refrozen cracks, which were numerous in this particular dataset, as can be seen from **Figure 4.1**. Maps of the difference between scaled surface elevation and draft were plotted and the AUV swath translated and rotated to find the optimum match. AUV swaths which overlapped each other were checked for consistency once they were matched to the laser data. Sonar data was discarded if the optimum position from the consistency check differed from the laser-sonar match. This dominantly occurred following the turn for home at the furthest point of each line (*i.e*. mid mission) and we preferred straight tracks for the composite dataset. We present data extracted from a total of eight runs (of 21 total). Multiple draft values from valid overlapping sonar swaths at any given location were averaged in the final mosaic. Over the critical area of the FY ridge, values of sonar draft (particularly the location of the deepest part of the ridge keel) were checked against a grid of 65 holes which were drilled in six lines across the ridge using Kovacs two-inch diameter drill flights. These were located on the laser master map with reference to the AUV deployment hut - a three-metre-high structure that was very obvious on the laser data - and to continuously-recording GPS base stations installed in the camp command hut and at the end of survey lines.

The matching process was iterative and time-consuming, but essential to give confidence in any comparison between measurement types. Data match to the nearest pixel (*i.e*. 1×1 m) at the tie-points. Between tie-points, it is likely - given the struggles with AUV navigation detailed in *Doble et al*. [2009] - that some mis-registration occurs. It is expected that such hard-to-quantify positioning errors have a small impact on the study, however, since it is hardest to achieve a match over exactly those relatively homogenous regions where any mis-registration will have the smallest effect. Consequently we feel justified in comparing the surface elevation and draft datasets at the highest resolution of 1×1 m, though the majority of analysis presented here in fact takes place at larger scales (typically 11 m diameter patches).

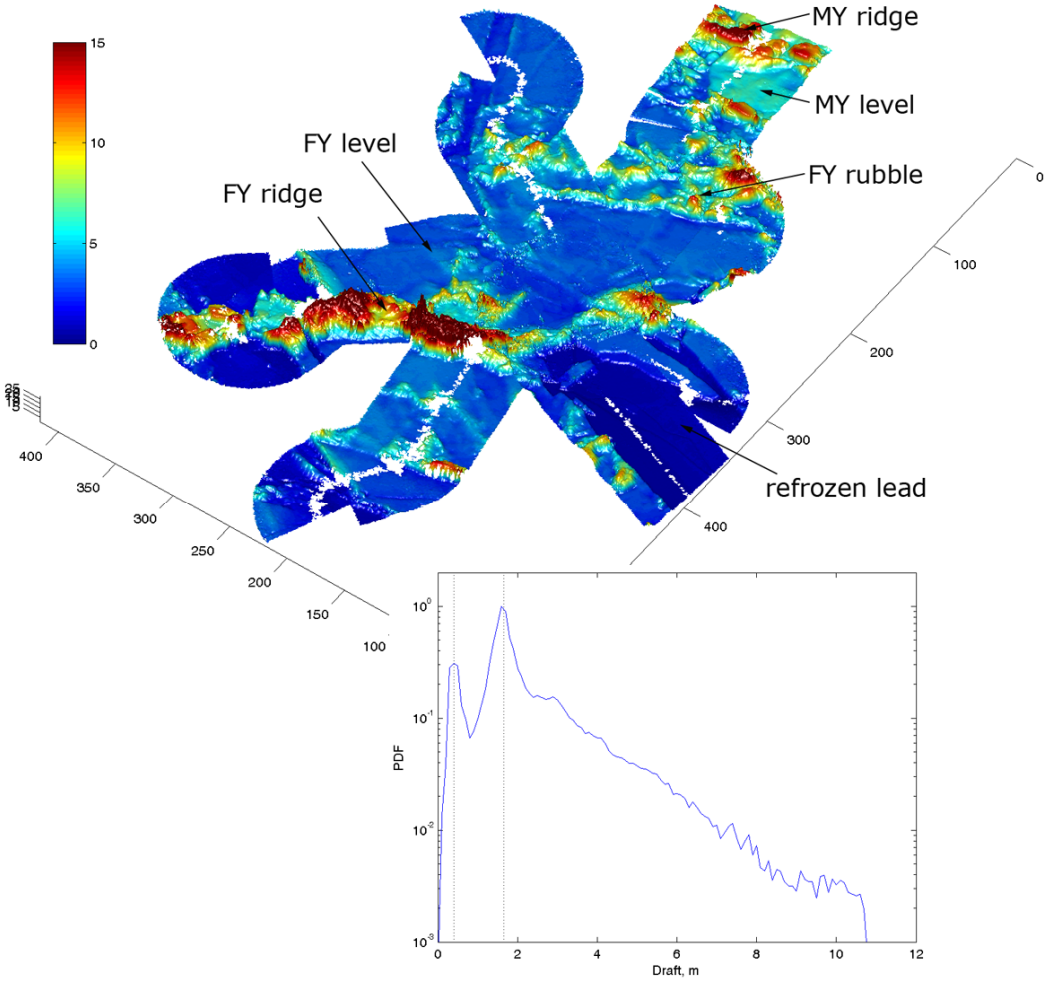

Figure 4.1: Three dimensional view of the ice draft mosaic from eight runs of the AUV-mounted inferometric sonar. Axes and colourbar are shown in metres and examples of the various ice types in the image are arrowed. Gaps are present from acoustic shadowing and the poor returns at nadir over level ice from the inferometric sonar system. The PDF of the AUV data at 1×1 m resolution (downscaled to the resolution of the co-incident laser surface elevation dataset) is shown inset. Modes at 0.40 m (refrozen leads) and 1.65 m (level FY ice) are marked as dotted lines.

4.2.4. In situ measurements of snow, ice and water properties

High technology survey instruments, as used in this study, do not obviate the need for detailed surface measurements to establish the physical parameters of the site. Critical parameters in determining the *R* value are the density and thickness of the snow and ice and the surface seawater density. Extensive measurements of these were therefore made during the ice camp. Surface seawater density (ρ_w) was calculated - from measurements made using a SeaBird 19 CTD lowered to a depth of 50 m - as 1021 kg m^{-3} (26 psu).

A grid of 143 holes was drilled on level FY ice (see Figure 1, *Wadhams & Doble*, 2008) to ground truth the AUV sonar data at its first deployment site, approximately 2 km from the data presented here. Snow thickness was measured at 3 points around each hole, giving 429 measurements in total.

Median snow thickness (h_s) was 11 ±9 (Standard Deviation) cm. Snow density (ρ_s) was 240 ±16 kg m^{-3}, from 12 snow samples distributed over the sampling area. Median ice draft (d) was 137 ±15cm, with an ice freeboard of 13 ±5 cm. From these measurements the *in situ*, or isostatic ice density (ρ_i) is 918 ±29 kg m^{-3}. The measured region was far (>300 m) from any deformed zones which might have disturbed its isostatic balance. Measured mean FY ice core density at the same site, from two side-by-side cores was 914 and 920 kg m^{-3}. This very close agreement suggests that the ice density measurement is reliable to better than 10 kg m^{-3} and is itself in very good agreement with isostasy, reflecting the very low porosity of the winter ice and the subsequent negligible drainage of fluids from the extracted core.

Long transects were performed using a Geonics EM-31 electromagnetic induction device, with snow depths determined every five metres with a ruler. Six radial lines, centred on the ice camp itself, were performed, out to a distance of approximately 1000 m. In the area of the co-incident measurements presented here (within 200 m of the ridge) the snow thickness was highly variable. Mean snow thickness was 27 cm, with a Standard Deviation of 29 cm. Maximum snow thickness was 1.19 m.

4.3. Comparing draft and surface elevation

We compare the surface elevations and drafts across the area covered by both laser and AUV measurements. AUV data were downscaled to the resolution of the laser data by taking the median value of the four 0.5×0.5 m AUV bins contributing to each 1×1 m laser bin. Total coverage of the co-incident dataset was 87,500 m^2 of ice.

4.3.1. Separating deformed and level ice

To understand the contributions to the observed top- and bottom surface relations, it is useful to consider the deformed and level ice fractions separately. Deformed ice regions were identified by examining the surface roughness of the sonar data. A circular patch of 15 m diameter, advancing in 1 m steps, was used to calculate the variogram (or semivariance) within the window [*Webster and Oliver*, 2007]. This patch size gives >100 bins per patch - a useful minimum figure to generate a representative variogram – while maintaining a reasonable resolution. In the sonar data presented here, variograms of deformed ice display peak magnitudes more than 100 times that of level ice, allowing them to be robustly segmented. The laser data is too noisy to classify in the same manner, however, and we use the sonar data to define "deformed" for both upper and lower surfaces. By this criterion, the co-incident dataset consists of 45**%** deformed ice by area, though this is likely to slightly over-estimate the true deformed ice fraction, since the edges of level floes are also classified as deformed ice.

4.3.2. The <surface elevation- draft> relation

Transforming the standard isostatic relation to relate ice draft, d, and ice+snow surface elevation, f (as measured by the laser) we obtain:

$$d = \frac{\rho_i f + h_s(\rho_s - \rho_i)}{\rho_w - \rho_i} \qquad \text{(Eq.4. 1)}$$

If snow thickness and the various densities are constant, this gives a straight-line relation between draft and surface elevation. At full resolution (1×1 m), however, the majority of <surface elevation – draft> pairings do not lie along the modelled relation: correlation between the parameters is effectively zero (r_S^2 = 0.07 – we quote Spearman's Rank correlation, r_S, rather than the more familiar r^2 since the data are not normally distributed, as is required for r^2 to be correctly applied). Most points lie on the 1.65 m modal draft value, reflecting the dominance of FY ice in the dataset, but these have a very wide spread of measured surface elevations, from 15 cm to over 1.4 m. A second horizontal band is seen at around d = 0.40 m, representing the thin, refrozen lead areas, which display a similarly wide range of measured surface elevations.

Averaging the draft and surface elevation over a larger area improves the correlation considerably. We form an approximately circular patch of a given diameter (an odd number of bins), and move this over the data (in both x and y directions) in half-radius steps, calculating the median value of draft and surface elevation within each patch. Though the value for a single patch is calculated as the median of bins within that patch, we only use bins within each patch which have been classified as level or deformed, according to the ice type under consideration.

Figure 4.2 shows the results, for level and deformed ice fractions. Results are shown as a contour plot of $\log_{10}$ (the number of points having a particular <surface elevation – draft> pairing) and mark the modelled relation given by Equation (4.1). For level ice (top panel), a fixed snow thickness of 30 cm is marked: this intercepts the crossing of modal values and is close to the *in situ* mean measured figure. The relatively small patch of only 11 m diameter – chosen to mimic the expected footprint of the future ICESat-2 laser - increases the correlation between draft and surface elevation to r_S^2 = 0.70, collapsing the previously extended range of surface elevation values at the modal draft and refrozen leads to a far smaller range. These dominant <surface elevation – draft> pairings now lie close to the modelled line, though surface elevations are seen to be positively biased with respect to this line, suggesting positive snow thickness anomalies from the 30 cm value which intercepts the most common values.

The contrast between full-resolution (1×1 m) and averaged (11×11 m) patches is explained by examining the histogram of the laser surface elevation data, which only displays a single mode at full resolution: the dominant level ice types are not resolved. Patch sizes of 11 m diameter and upwards split the laser surface elevation histogram into two distinct modes, differentiating the refrozen lead (mode at 26 cm) from level FY ice (mode at 42 cm). We note that these two peaks are separated by almost exactly the stated RMS error of the laser instrument. Clearly the relatively high noise of the laser instrument requires averaging if sufficient signal:noise ratio is to be obtained over such areas. The sonar drafts easily resolve the same ice types without further averaging, since (a) the accuracy has already been significantly increased by combining individual range-angle pairs into bins, inside the proprietary *GeoSwath* software; (b) variations in the draft values are more pronounced than surface elevation, in line with isostasy.

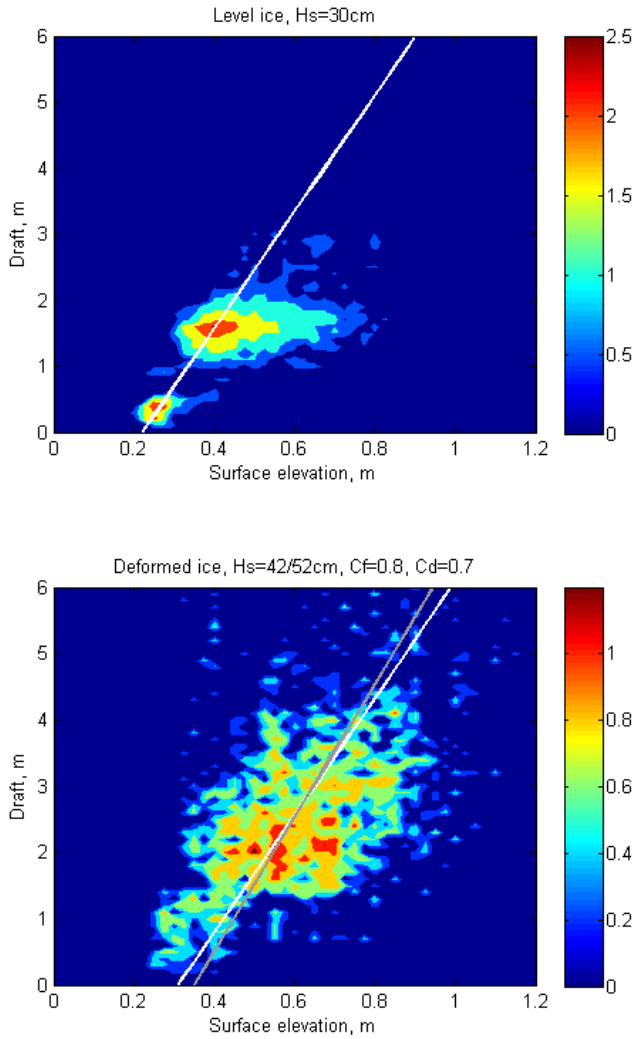

Figure 4.2: Contour plot of measured surface elevation against draft, for level ice (top panel) and deformed ice (bottom panel). Data are medians over 11 m diameter patches in both cases. The colour scale shows the base 10 logarithm of the number of bins occupying each <surface elevation – draft> pairing. The inclined white line shows the modelled relation, given by Equation (1), for the indicated fixed snow thicknesses. The deformed ice plot also shows the relation (grey) taking into account typical porosities for above- and below-water deformed ice features.

For the deformed ice fraction (**Figure 4.2**, bottom plot) we choose a fixed snow thickness which minimises the total error to the fit line, defining $\overline{\Delta d}$, the mean draft error from the (isostatic) fit line:

$$\overline{\Delta d} = \frac{1}{N}\sum_{f}\sum_{d}\Delta d.n_{f,d} \qquad \text{(Eq. 4.2)}$$

where $n_{f,d}$ is the number of bins having any given value of surface elevation (f) and draft (d), Δd is the difference between the draft value under consideration and the draft value of the fit line at that

surface elevation, and *N* is the total number of patches. For the 11×11 m data shown, this gives zero error at 42 cm snow thickness (plotted as the white line on **Figure 4.2** bottom plot). For the deformed ice fraction the effect of porosity was also investigated, taking account of standard solid ice fractions above water (C_f = 0.8) and below water (C_d = 0.7) [*Lepparanta and Hakala*, 1992; *Melling et al.*, 1993]. Taking porosities into account, the zero $\overline{\Delta d}$ value was achieved at 52 cm snow thickness, plotted as the grey line.

The data form a much more diffuse 'cloud' for deformed ice than for level ice. Since we have fitted the optimum snow thickness, the points lie symmetrically around the fit line, though we note that at low drafts the data are biased to a reduced snow thickness (to the left of the fit line). Evidently, 42 or 52 cm snow thickness is unlikely to exist on ice whose ice+snow surface elevation is less than these figures, in contrast to the Antarctic where negative ice freeboards are common.

The spatial pattern of the major departures from point-to-point isostacy are shown in **Figure 4.3,** which plots *R* over the co-incident region at full resolution (1 m bins). It is clear that the highest *R* values (*c*. 10) are associated with the FY ridge, where the sonar data show a much wider keel than the relatively thin sail seen by the laser. High values are also observed in other deformed ice areas, such as the FY rubble field and in the MY ridge. Conversely, the lowest values (1-2) are observed in the thinnest ice (refrozen leads), reflecting the relatively high contribution of snow thickness to the surface elevation there. The modal value of *R* over the whole dataset is 3.4, with the level FY ice areas dominating this contribution. The inset cumulative distribution function (CDF) shows that approximately 50% of the data lie more than *R*±1 from the mode (*R* = 2.4 intercepts the curve at 22%, *R* = 4.4 at 72%). Such large departures from the level ice *R* clearly show that much of the region is not in the point-to-point isostatic balance that homogenous material properties would suggest. We postulate that the high *R* values lie dominantly in the deformed ice areas and examine the variation of *R* with draft in the next section.

4.3.3. R versus *draft*

The relation between draft and *R* is displayed in **Figure 4.4**, again split into level ice (top plot) and deformed ice (bottom plot) fractions and calculated taking median values for 11 m diameter patches. The colour scale again shows the base 10 logarithm of the number of bins, this time with any particular <draft – *R*> pairing. As previously suggested by **Figure 4.2**, the majority of level ice points are clustered around the FY level ice draft value of 1.65 m, with a large variation in *R*: a significant number of points are encountered with ratios between 2 and 5 for this draft. Also as seen in **Figure 4.2**, a second core of values is clustered around 40 cm draft (refrozen leads), with *R* values of around 1.5. Evident in this plot is a general increase of *R* with draft.

The white curve shows the expected relation between draft and *R*, eliminating ice freeboard from the isostatic relation (Eq. 4.1) and expressing it in terms of densities, draft and snow thickness only:

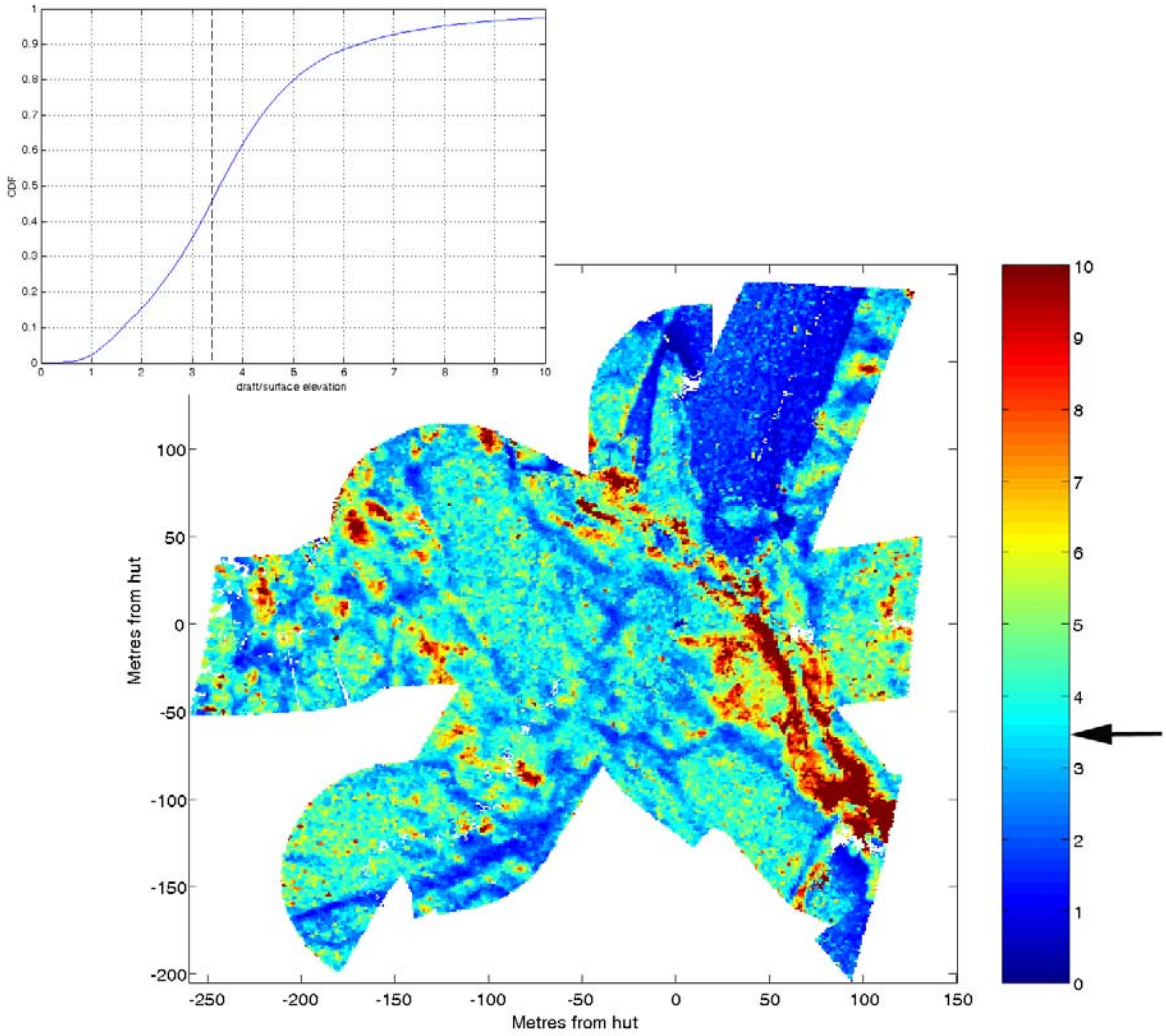

Figure 4.3: Map of *R* calculated at full resolution for all valid bins. The arrow next to the colour scale marks the dominant (modal) level ice value of *R* = 3.4. Co-ordinates are referenced to the AUV deployment hole and hut. A relation with ice features is clear, with high values (~10) over ridge features and low values (1-2) over the thinnest ice, in refrozen leads. The inset curve shows the cumulative distribution function of the plotted *R* values, with the level ice value marked as a dashed line.

$$R = \frac{\rho_i d}{d(\rho_w - \rho_i) + h_s(\rho_i - \rho_s)} \qquad \text{(Eq. 4.3)}$$

Fixed values are used for ice density (918 kg m^{-3}), snow density (240 kg m^{-3} – note that this is rather lower than the typical values given by the Warren climatology and used by the AWI Level-1b processor) and snow thickness, as before. The contribution of (a) snow thickness variability and (b) laser measurement error to the observed variability of *R* was then assessed. For each draft, the snow thickness was varied and Equation (4.3) recalculated, to give an isostatically compensated error. Results are shown as the light blue curves. Snow thicknesses between 20 and 70 cm (top panel) cover most of the observed variability (3< *R* <6). The effect of the laser measurement error of 15 cm RMS

per bin was assessed by adding/subtracting that error without recalculating the isostatic relation. If the error is independent between bins, then the error for each patch is reduced by the square root of the number of bins making up the patch (73 for an 11-bin diameter patch), giving 1.8 cm error in this case. The yellow curves plotted on the top panel indicate that such an error is rather small, though in fact the error is unlikely to be completely independent from bin to bin.

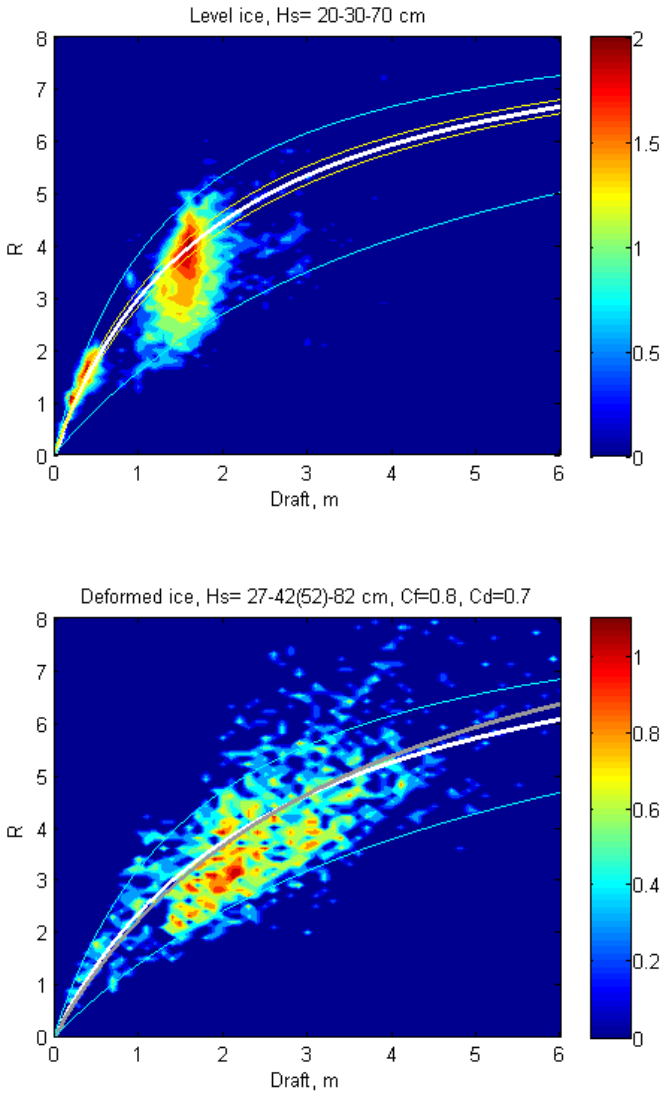

Figure 4.4: Contour plot of *R* against draft, compared using the same 11 m diameter patches used for Figure 4.2. The colour scale shows the base 10 logarithm of the number of bins contributing at each *d*-*R* pair. Top panel shows the result for level ice, bottom panel for the deformed ice fraction. Also plotted is the modelled relation using appropriate densities (white curve), with a fixed snow thickness of 30 cm (top) and 42 cm (bottom, white curve) and with typical ice porosities and 52 cm snow thickness (bottom, gray curve). Yellow curves show the effect of the expected laser freeboard measurement error (not isostatically compensated). Light blue curves show the effect of isostatically-compensated snow thickness variability, for snow thicknesses between 20-70 cm (top panel) and 27-82 cm (bottom panel).

Figure 4.4 (bottom panel) shows the <draft – *R*> relation for deformed ice, also using 11 m diameter patches. The data, especially in the core region centred on 2 m draft, are much better aligned with the modelled curve, which uses the optimum fixed snow thickness of 42 cm (no porosity, white curve) and 52 cm (grey curve, using typical porosity values above and below water). 17% of patches have *R* values more than twice that seen over level ice. The blue curves plot isostatically-compensated snow thickness between 27 and 80 cm, which include the majority of data. There are a significant number of points at *R* >6, however, which may be attributed either to a reduced snow load over deformed ice, reflecting the often bare-ice nature of ridge crests; to a decreased effective density of the ice-snow composite in deformed features, as evoked by the inverse thickness-density relation of *Ackley et al*. [1976]; or – probably dominantly – due to the fact that deformed features tend to be much more extensive underwater than above.

4.3.4. FY ridge

We investigate this last issue – the extensive nature of underwater deformed features and therefore the difficulty in estimating their contribution from surface measurements – by isolating and examining the FY pressure ridge, comparing the total underwater ice volume that would be predicted by isostacy from the laser data with the volume actually measured by the AUV sonar.

The co-incident dataset was cut down to an area of 76×136 m around the ridge and the ridge further delimited by removing points less than a given draft threshold. We calculate the measured *R* value over the thresholded region as the sum of drafts in the 1×1 m bins (*i.e*. ice volume underwater), divided by the total volume in the same region as seen by the laser (*i.e*. the ice volume above water). This is equivalent to the mean *R* value over the ridge. We compare this figure to the *R* value predicted from the surface elevation by isostacy, according to Equation (4.1), for a range of threshold drafts, in **Figure 4.5**.

Examining the thresholded regions by eye suggests that a threshold of 3.0 m is most appropriate in this case – approximately double the surrounding level ice thickness of 1.65 m. Over the whole ridge area, the measured *R* value is 7.7 at this threshold, while isostacy would predict a value of 6.6. The discrepancy increases with greater thresholds, as the slopes leading up to the keel maximum are excluded from the below-water calculation, while the significant ridge sail (which is centred above the keel in this case) remains to contribute to the above-water component. Thus, the isostatic relation tends to significantly underestimate the actual draft in the ridge area. Taking representative porosities for the ridge keel (30%) and sail (20%) into account shifts the overall *R* values for the ridge to 6.5 (measured) and 5.8 (predicted), at a threshold of 3.0 m.

4.3.5. Regionally-averaged R

It is clear that the size of the patch within which we calculate statistical values plays an important part in constraining the measured values to a workable <surface elevation → thickness> relation, trading spatial resolution for accuracy in ice thickness. We therefore examine how the descriptive statistics change as the averaging scale (footprint) is increased.

The median within circular patches is therefore calculated, from full 1×1 m resolution to 300 m diameter patches encompassing the whole dataset - similar in area to the footprint of CryoSat-2's radar altimeter in sea ice mode [*Wingham et al.*, 2006]. As before, a circular patch of a given diameter is run over the data at half-radius increments in both *x* and *y* directions. For surface elevation, we also calculate the result using the minimum value within each patch, as recent work suggests that CryoSat retrievals are largely determined by the thinnest ice within a footprint [*Tonboe*

et al., 2010]. ICESat is assumed to return the mean surface elevation within a footprint over snow-covered surfaces [*Kwok et al.*, 2007] and we calculate the mean within each patch accordingly. We also calculate the mode (both within a patch and mode of all modal values). In each case the *R* value is calculated using the median draft value within the corresponding patch, since comparing methods with this dataset, it was found that the median gives the most stable measure of *R* at any scale. **Figure 4.6** presents the results, again split into level ice (top panel) and deformed ice (bottom panel) fractions.

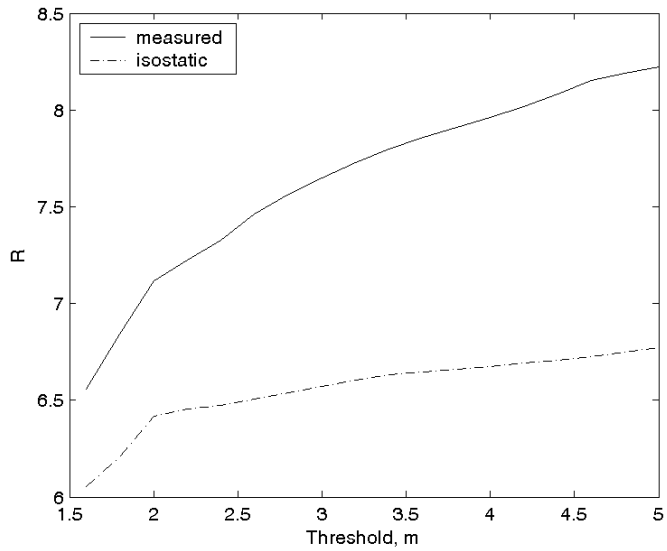

Figure 4.5: *R* values over the FY pressure ridge, calculated as 'measured' (mean draft / mean surface elevation) and 'predicted' (draft calculated from surface elevation by isostacy), for various threshold values of draft used to define the ridge keel.

For level ice, the mean and median relations hardly deviate from their full-resolution value of *R* = 3.4 at any scale, reflecting the dominant nature of level FY ice - with its relatively tightly constrained *R* values - in this dataset. The ICESat-like response (mean surface elevation) is particularly stable, deviating by less than ±0.1 over the whole range, while the median rises slightly to 3.6 at large scales. The inter-quartile range for this solution (plotted in grey) more than halves over the same scale range (25th and 75th percentile ranges marked). Modal values diverge to *R* = 4.1 at the full dataset size. The <median draft – minimum surface elevation> curve (not plotted) quickly outstrips reasonable *R* values, reaching *R* = 43 for the largest patch. Since a single 1×1 m bin is unlikely to greatly perturb CryoSat returns, we instead calculate the surface elevation of the lowest 10% and lowest 25% of bins within a patch. *R* values for both solutions are significantly higher than either the mean or median results, reaching *R* = 4.3 and 5.2, respectively, at the largest scale.

At averaging scales of 11 m and above, the mode and mean diverge from the median value - the mode increasing while the mean drops. Examining probability density functions of the *R* population at various scales shows the reason for this: Larger footprints allow the laser to resolve thin ice patches, splitting the distribution of *R* values into two peaks and eliminating *R* values between 1.8 and 2.8 from the statistics. For level ice, the scale dependence of *R* is thus largely dependent on the

characteristics of the laser measurement, at least for these data which are dominated by the *R* values for two ice types - FY level ice and thin, refrozen leads.

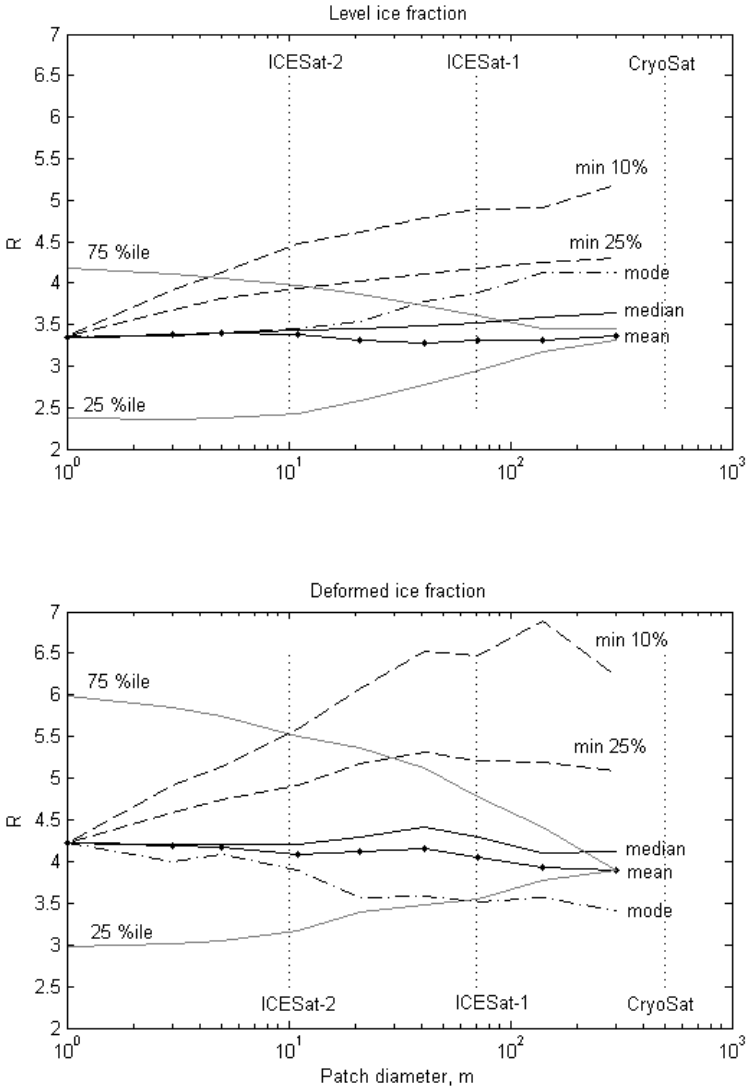

Figure 4.6: Regionally averaged *R* values, for level ice (top panel) and deformed ice (bottom panel), over circular patches ranging from 1 – 501 m diameter. Surface elevation values are calculated as (a) the mode; (b) the mean; (c) the median; (d) the minimum 10%; (e) the minimum 25%, of values within a patch. Draft values are calculated as the median of bins within a patch in all cases. *R* values are then calculated as the median of all patch values, reducing the data to a single value for each scale. 25% and 75% interquartile ranges are marked for the mean solution (dashed lines). Typical ICESat-2, ICESat-1 and Cryosat footprints are indicated as vertical dotted lines.

The deformed ice fraction (**Figure 4.6**, lower panel), exhibits more variability, as would be expected. The mean and median solutions exhibit similar behaviour, being rather stable around the full resolution value of 4.2, with the mean tracking slightly lower than the median. Interquartile range is significantly larger than for level ice, at $3 < R < 6$ at full resolution (*c.f.* $2.4 < R < 4.2$ for level ice).

In contrast to level ice, the modal solution is lower than either the mean or median, dropping to 3.4 at the largest scale: the probability density function becomes more peaked, losing data on the upper tail as the mode shifts to lower *R*s. No secondary peak emerges due to laser resolution issues, in contrast to the level ice fraction. The minimum 10% and 25% solutions are predictably much less stable than for level ice, rising to maxima of 6.9 and 5.3, respectively.

4.4. Comparison with Cryosat results

We can compare the *R*/draft plots from these *in situ* experiments with that generated in the same region (Beaufort Sea) at the same time of year (April), though since CryoSat was not operating in 2007, we take the first available season, April 2011. The results are plotted in **Figure 4.7**. Note that we add back in the penetration depth, so these results are calculated for the same "top of snow" surface elevation as for the laser *in situ* results.

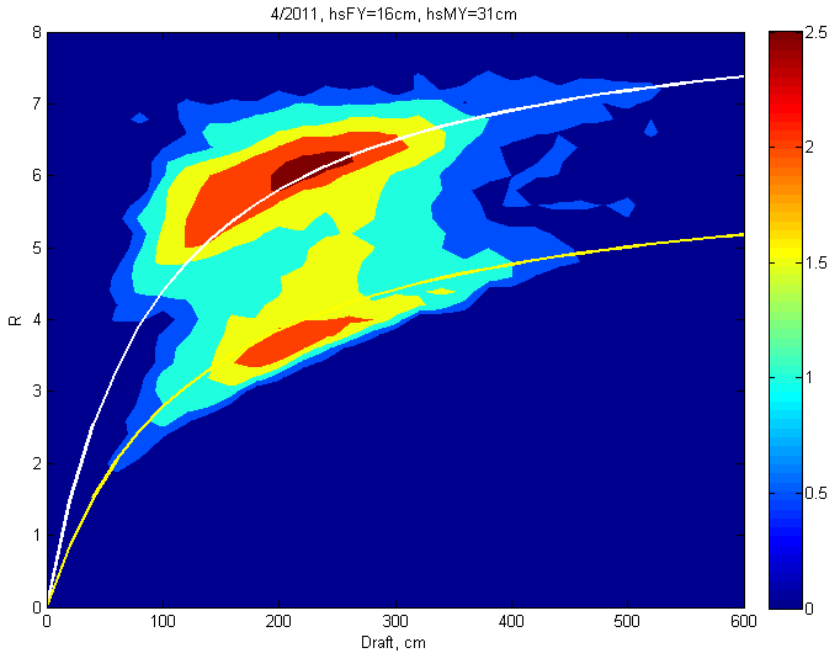

Figure 4.7: An equivalent plot, generated using the AWI CryoSat product for April 2011. The coloured contours show prevalence of any particular *R*/draft relation, as before, together with the theoretical curves for the imposed snow thicknesses of 16 cm on FY ice (white curve) and 32 cm on MY ice (yellow curve). The upper 'blob' represents the FY ice area, with a typical *R* of around 6, the lower blob is the MY ice region, displaying an an *R* value of around 3.5, due to the increased snow thickness beyond the radar penetration depth and lower ice density.

The calculated *R* values plot neatly against the expected curves for fixed snow depths of 16cm (white curve) on FY ice and 32 cm (yellow curve) on MY ice. Clearly the core region (R=5.8, draft=2.2m) has values rather different from those measured in 2007 (*e.g*. **Figure 4.4**).

The *R* values plotted above are in fact rather typical of that derived by the AWI processor at all available months. **Figure 4.8** shows the derived *R* value for the 27 available months, for the camp location (73N, 145W) FY ice (top panel), and for a dominantly MY ice cover, near the northern limit of the instrument, at 87N,145W (bottom panel).

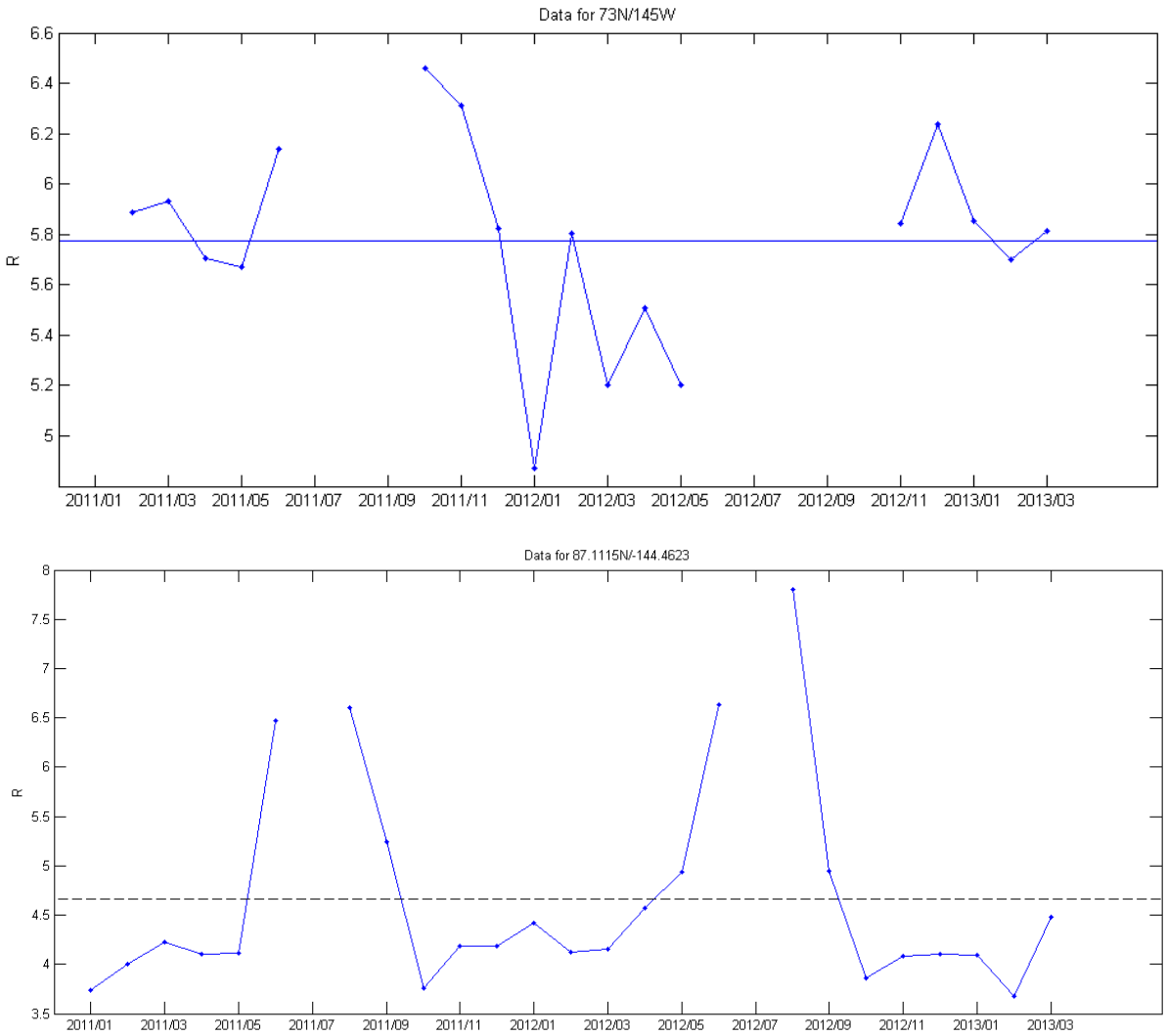

Figure 4.8: Calculated Cryosat *R* values for all available months, at the ice camp location (73N, 145W, top panel) and at a dominantly MY site (87N/145W, bottom panel). The mean value is indicated as a horizontal line in both plots.

The Cryosat-derived *R* at the camp location (*R*=6) is clearly very different than that seen in the Gavia/laser data for our dominantly FY ice region (*R*=3.4). The MY ice region (R~=4) matches the in situ measurements more closely, in accordance with the more closely-matched snow depths.

The widespread nature of these characteristics is evident if we plot the Arctic-wide *R* for a given month, as shown by **Figure 4.9**, here for March 2012: *R* values divide clearly between FY ice (red, *R*~=6) and MY ice (blue, *R*~=3.5). The contrast with *R* plotted over the measured area (**Figure 4.3**) is very marked. 2012 itself is typical of all three years, as shown by the histograms of Arctic-wide *R* values for March 2011, 2012 and 2013 (**Figure 4.10**).

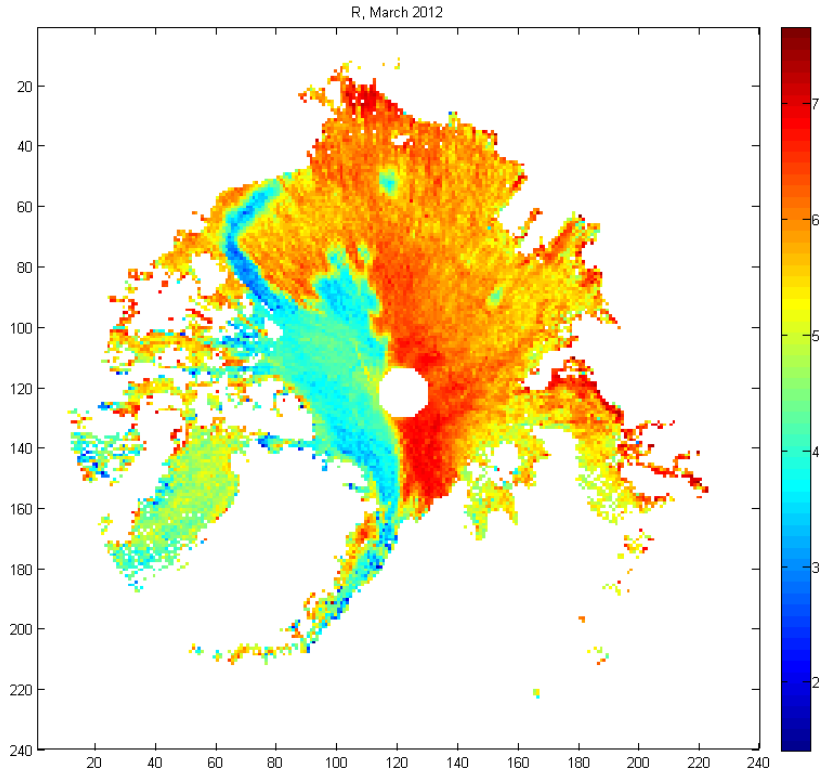

Figure 4.9: Calculated R values over the whole Arctic for March 2012.

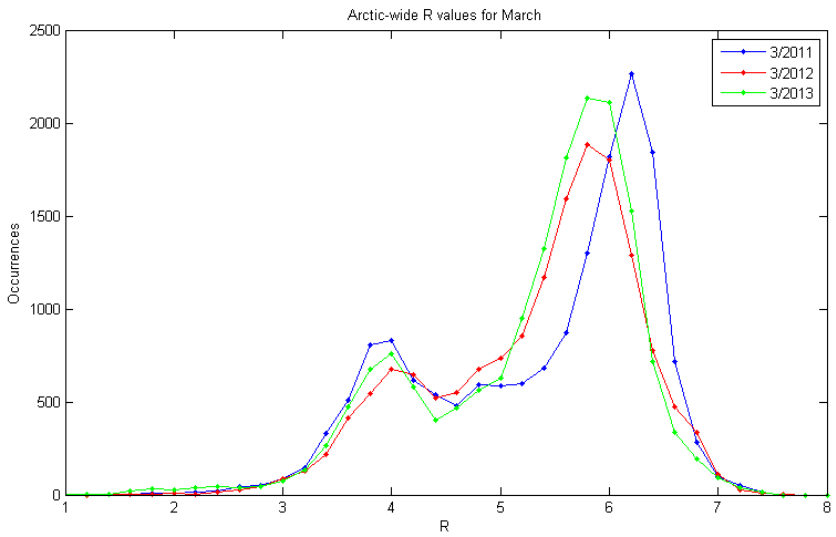

Figure 4.10: Distribution of *R* values across the whole Arctic for March 2011, 2012 and 2013.

Figure 4.11, below shows graphically the difference between the CryoSat result and the *in situ* measurements. Though 2007 may well have been different to the 2011 CryoSat result, our measured FY ice thickness values are very similar to that seen in the BGEP mooring data, suggesting that this end-of-winter FY ice thickness is rather typical of the region.

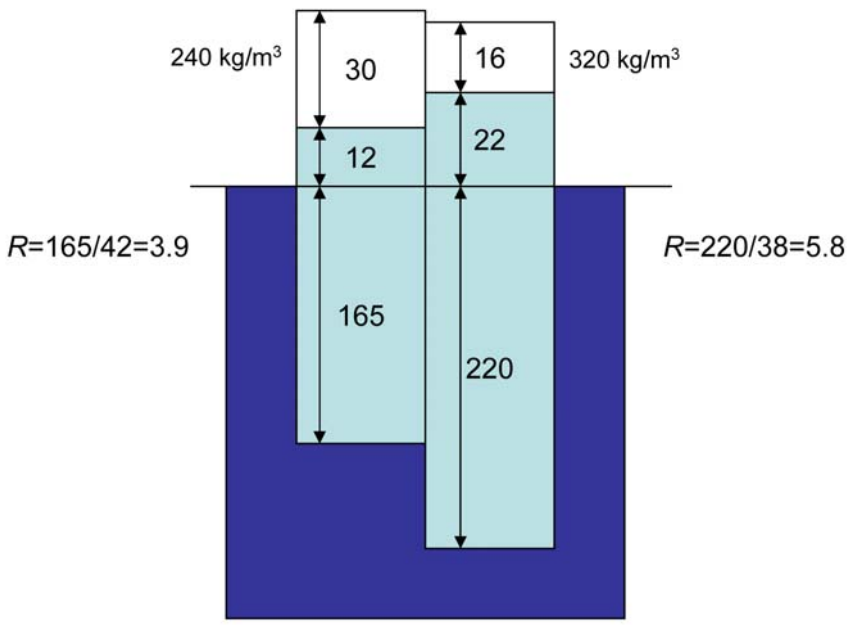

Figure 4.11: In situ measurments (LHS) compared with measured and assumed CryoSat data (RHS) for the same point and month.

The discrepancy over the relatively simple case of uniform level ice lies in the snow depth variation, since it is clear that this is highly variable over footprint–sized scales. Kurtz and Farrell (2011) analysed the Ice Bridge snow radar data for 2009 and suggest that 16 cm is indeed a realistic figure for snow depth over FY ice (33 cm over MY ice). Had this been the case during the APLIS survey analysed here, then the measured draft/freeboard ratio would reduce to 165/28 = 5.8, in line with the CryoSat result.

4.5. Discussion and conclusions

There is a general increase in *R* with ice thickness, as noted by previous investigators – **the opposite to that modelled by the CryoSat processor** - and a strong tendency for the data to display lower *R* values than the modelled curve over level ice. This can be attributed either to a dominantly positive snow thickness anomaly or, equivalently, a lower composite material density than the measured fixed ice density would suggest. No evidence of thickness-dependent material density could be found, however, in contrast to earlier work by *Ackley et al*. [1976] and *Kovacs* [1997]. For the current dataset, material densities present a wide range of values at the two dominant drafts, which

remained true whether the values were calculated at full resolution or over larger patch sizes. This finding agrees with the conclusions of *Wadhams et al*. [1992] and *Forstrom et al*. [2011] whose calculated material densities were consistent over their study regions.

Over deformed ice, the apparent material density was in fact higher, opposite to that suggested by *Ackley et al*. [1976] and *Kovacs* [1997]. This can be attributed both to the fact that deformed ice features tend to be wider underwater than above, as well as to the highly porous nature of the young FY deformed ice, where seawater in the open pore spaces significantly increases the overall features' apparent density when its envelope is measured, as done by the sonar. These factors result in the **highest *R* values occurring over deformed ice**. The <*R* – draft> relation follows the modelled isostatic curve more closely over deformed ice than for the complete dataset, since variations in *R* are more closely tied to ice thickness over deformed ice, as opposed to the snow thickness variability which apparently dominates the variability of *R* over level ice in these data.

Examining how *R* changes with footprint size and averaging method shows it to be rather stable at over all scales examined, from full-resolution 1×1 m bins to a patch size covering the complete dataset, at least where the mean and median of values within each patch are used. For level ice, the dominance of FY level ice in these data accounts for this: this ice type is sufficiently extensive and homogenous that it matters little what footprint size is used. The result is perhaps more surprising for deformed ice, whose magnitude of variation of mean and median with scale is very similar to that seen for level ice (*c*. 0.3), though the interquartile range is nearly three times the spread of level ice values: the PDF of deformed ice *R* values becomes more peaked with increasing scale, but does not change in its essential character. The PDFs of *R* for level ice change considerably, however, due to the averaged laser measurements being able to resolve the thin ice peak at patch sizes above 11 m, and results in the level ice mode evolving in the opposite sense (*i.e*. increasing) to the deformed ice fraction.

The CryoSat simulation – taking a surface elevation corresponding to the minimum 10% or 25% of values within a patch – is less successful, significantly over-estimating *R* at representative footprint sizes (e.g. *R* >6 for minimum 10% values over deformed ice and *R* >5 over level ice). The mechanism by which the thinnest ice dominates the CryoSat return is not a simple 'percentage minimum' as used in this simulation, however, but is dependent on surface roughness: "*The high backscatter magnitude from the thinnest ice within the footprint largely determines the elevation of the effective scattering surface*" [*Tonboe et al*., 2010] and has a significant effect down to very low percentage area coverage of the thinner ice type. Though it should be borne in mind that the CryoSat radar penetrates the snow layer, unlike the laser measurements considered here, it is clear that such characteristics will significantly affect CryoSat's ability to correctly report ice thickness over the varied ice terrain which forms the vast majority of Arctic pack ice.

The CryoSat *R* values over FY ice (**Figure 4.9** and **4.10**) are radically different to that seen in these *in situ* data, and suggest that the assumed values of snow thickness, snow density, snow penetration, ice density and water density are all adding errors to make the final derived ice thickness significantly different to that measured *in situ*. The **over estimation of ice thickness is clear for FY level ice** in both this Section's data and from the BGEP mooring data comparison in Section 3.

CryoSat's *R* value over deformed ice is much closer to that observed here (3.5 *c.f*. **4.2**). This arises from; (a) higher assumed snow thickness over MY ice; (b) a lower fixed ice density over MY ice (882 kg/m^3 *c.f*. 916 kg/m^3). Of course, not all deformed ice is MY ice.

In situ data from deformed ice regions has shown that CryoSat results are more unreliable, generally under-estimating the ice thickness (see Section 3, EM and Ice Bridge comparisons). **Figure 3.7** (red and blue curves) strongly suggests that **the CryoSat return is indeed coming from the thinnest ice in**

the footprint – or at least the thinnest ice which isn't cracked/contains open water in the footprint – as has been suggested by Tonboe et al. (2010), though Ice Bridge data is more equivocal.

5. Estimating the regional importance of deformed ice, using automated classification of manned submarine upward-looking sonar data

5.1. Introduction

The previously-stated investigation of the freeboard-draft relation for level and deformed ice suggests that the CryoSat returns may be particularly problematic over deformed ice. We therefore move on to estimate the prevalence of deformed ice in various regions of the Arctic, and hence determine where the CryoSat results are likely to be the most problematic.

To this end, we examine upward-looking multibeam sonar data, collected during a March 2007 transect of the Arctic Ocean by the UK submarine *HMS Tireless* – the first time that multibeam had been fitted to a long-range under-ice vehicle. Ice draft measurements are calculated from the ranges and angles of each beam, and combined with the depth, roll, pitch and heading of the submarine. The post processing of these data was carried out using MB software (an on-line freeware system designed to handle all kinds of multibeam sonar data). Problems with the sonar itself (a water leak into the cables connecting the sonar head to the control unit) and with the continuing unavailability of the submarine attitude/navigation data, have forced us to be selective in the choice of data. We here analyse and report on a total of 285km of data, from three regions: Fram Strait; North of Greenland; and the Beaufort Sea.

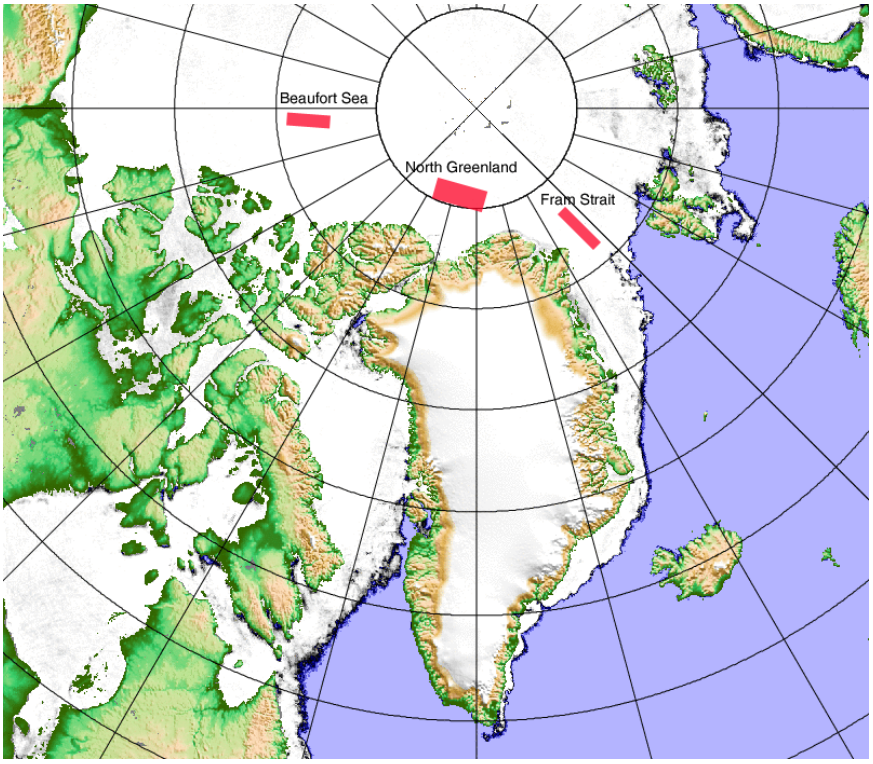

Figure 5.1: Map of the three regions discussed in this section, overlaid on the satellite-derived ice cover for 12 March 2007 (the approximate date of the survey).

We developed an novel ice type classification algorithm, using only the surface roughness properties of the ice underside to classify the ice into one of six classes: four classes of level ice – 'unknown age', FY, 'mixed age' or MY; and two deformed ice classes – 'deformed', or rubble; and ridged ice.

5.2. Multi beam sonar data description

The *Tireless* dataset was received as ASCII files containing data values for ice draft, backscatter and submarine keel depth for each along-track instance (ping) and cross-track beam (160 beams). Each ping/beam point was associated with a latitude and longitude. The individual ping/beam data are rather data dense (around 12 MB for 1000 pings) and the next stage is to reduce this by binning the values into a regular grid. We choose a 2 m bin size, which is the best compromise between resolution (higher resolution with smaller bins) and achieving a reasonable coverage of the swath (bigger bins mean less gaps), for the current dataset where the submarine is running between 80 and 115 m depth. Binning also serves to reduce error from individual ping/beam values, since multiple values inside a single bin are averaged together in a weighted mean scheme. We track how many points have contributed to each bin, a number which ranges between 3 and 40 on average.

Values are clamped to a limited width either side of the swath centreline, since the outermost beams generally give poor quality data. Erroneous zero draft values, associated with displaced navigation data, are also removed: any single ice draft value that differs from its neighbours by more than 3 m is classed as a spike and removed. The native latitude/longitude reference system was converted to an x/y (metres) reference system and the lat/lons discarded in accordance with Royal Navy security instructions.

We perform a basic correction for boat depth variations. The submarine typically 'porpoises' underwater, with boat depth varying by about 1 m on a horizontal length scale of 100 m. In the absence of complete navigation data, discussed above, this variation is imposed on the ice draft. We thus filter boat depth for the central beam with a 200 m scale low-pass filter, and subtract the anomaly (difference from the mean boat depth) of this filtered output from the ice draft. This improves the appearance of the multibeam data only if the submarine is not making significant depth changes (*i.e*. is merely porpoising). If significant boat depth changes occur during a swath, we retain the un-corrected ice draft data. While this porpoising would be problematic if we were primarily interested in accurate ice draft values, it should not overly affect the surface-roughness-based classification that is the object of this study, since the length scale of the porpoising far exceeds that of the ice draft variations which we use for classification.

The binned files contain many data gaps, due to poor installation of the sonar and resulting water leaks into the connecting cables between sonar head and processing unit. These 'gappy' files make it difficult to see what structures or ice types are present. Swaths were therefore interpolated to guide the eye when evaluating classification routines. We emphasise that the classification takes no account of this interpolated data, operating purely on the original data values. Interpolation is done using the tri-scattered method, using natural neighbours. **Figure 5.2** shows the results of interpolation on an example line.

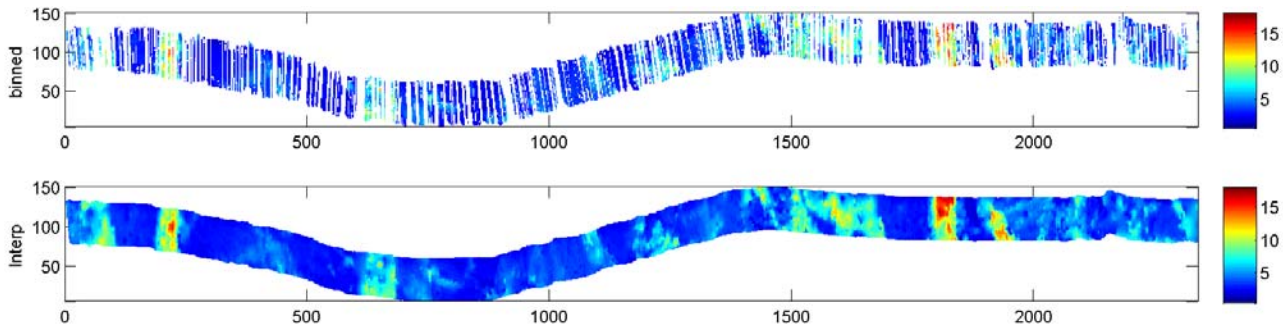

Figure 5.2: Binned (top panel) and interpolated (bottom panel) ice draft, for swath mb0027. Colour scale is 0-15 m ice draft in both plots.

The various factors reducing data quality also reduced the total track length of data that we were able to produce. We were able to retrieve 55 swaths in all, with a total along-track length of 285 km. Data are drawn from three diverse regions in the Arctic. **Table 5.1** gives a summary of the swaths and their properties.

- Fram Strait: 12 swaths, in what we expect to be a mixed FY and MY ice cover
- North of Greenland / DAMOCLES fuel cache area: 37 swaths, in heavily deformed, dominantly MY ice
- Beaufort Sea: 6 swaths, in dominantly FY ice

Table 5.1: Summary of imported data. Swaths are grouped by region. "Pings" gives the number of along-track pings (unbinned data) making up the swath; "Bins" gives the number of along-track 2 m bins in the binned swath; "Mean/mode" give the mean and mode(s) of the binned ice draft. "Pop" gives the mean number of individual ping/beam points contributing to a bin. Total along-track length (km) is calculated as the sum of along-track bins, x2 (since the bins are 2m across).

Filename	Pings	Depth, m	Bins	Mean, m	Mode(s)	Pop
Beaufort Sea						
mb0001	1001	83	888	3.3	2.9/2.6	8
mb1002	1539	74-68	645	3.2	2.3	12
mb1003	2342	67	836	3.8	2.4/2.3	15
mb1004	1516	80-83	739	2.4	2.8	7
mb1006	1259	53	162	3.0	2.6/2.6	40
mb1007	1111	83	749	4.4	3.9	9
North Greenland						
mb0009	5677	55	828	4.6	3.1/2.9	26
mb0011	433	112	382	5.3	4.1/-	5
mb0012	982	113	930	4.4	2.8/2.9	4

mb0013	2894	112	2833	5.1	2.7/2.3	5
mb0014	4102	113	3712	4.7	2.7/2.7	5
mb0015	1413	113	1182	4.8	2.7/2.5	6
mb0016	4083	113	1844	4.4	2.7/2.7	6
mb0017	4085	113	2734	4.5	2.6/2.3	6
mb0018	505	113	365	3.4	2.8/2.5	6
mb0020	4033	113	3676	4.5	2.5/2.7	5
mb0021	1329	113	1237	3.8	2.4/2.3	5
mb0023	3935	113	3763	4.5	2.9/2.6	5
mb0024	1241	113	1214	4.1	3.2/2.6	5
mb0026	3951	114	3755	5.3	3.0/2.9	5
mb0027	1194	114	1170	4.7	2.9/2.5	5
mb0029	3919	113	3771	5.2	2.9/3.0	5
mb0030	1157	113	1119	4.6	3.3/2.9	5
mb0032	3936	113	3673	5.4	3.2/2.9	5
mb0033	1660	113	1632	4.6	3.0/3.1	5
mb0035	3906	113	3694	4.5	3.3/3.1	5
mb0036	1854	113	1795	5.6	3.1/3.8	5
mb0037	2185	113	2121	6.1	3.3/3.0	5
mb0038	1015	113	985	6.1	3.3/3.0	5
mb0039	3804	113	2155	4.8	2.7/2.4	5
mb0040	2766	113	2737	5.4	3.4/3.2	5
mb0041	695	112	606	4.3	2.6/3.0	5
mb0042	3763	113	3665	4.4	3.2/3.2	5
mb0043	1365	113	1346	4.7	3.1/3.3	5
mb0045	2300	113	2311	5.5	3.0/2.9	5
mb0047	4882	55	307	5.3	3.2/3.0	30
mb0048	5125	58	1028	5.2	2.3/	33
mb0049	1604	70	453	5.2	3.4/	22
mb0051	1534	113	759	5.5	3.0/2.2	5
mb0052	2077	113	2346	4.4	2.7/3.1	5
mb0054	3179	113	3858	4.8	3.0/2.7	5
mb0055	1109	113	1310	4.4	2.9/2.5	5

mb0060	1300	83	461	4.7	2.7/2.4	12
Fram Strait						
mb2002	3886	142	6857	3.5	2.7/2.5	3
mb2003	5769	82	1495	4.4	2.1/2.1	13
mb2004	3358	83	750	2.0	0.6/	15
mb2005	3812	142	7650	3.5	2.7/2.5	3
mb2006	3980	142	7925	1.9	1.1/0.9	3
mb2007	3988	142	7782	2.0	1.6/1.6	3
mb2008	3850	143	7795	2.5	1.7/1.6	3
mb2009	3983	143	7844	3.0	2.6/2.7	3
mb2010	3859	143	7683	3.0	2.6/2.6	3
mb2011	3930	143	7559	2.9	2.5/2.4	3
mb2012	1500	143	3133	2.8	2.6/2.5	3
mb2013	1543	83	574	3.7	1.2/1.3	11
	Total track length, km		**285.65**			

5.3. Surface roughness classification method

Surface roughness parameters can be defined over an entire swath, but this gives little detailed information and mixes all the ice types present in the swath together. We must choose an area of investigation – a window – which is small enough to give a good spatial resolution to the classification, but large enough to contain enough data points (bins) to have robust statistical properties.

We examined the data with circular windows of various sizes. Circular windows of 15 bins diameter and greater gave stable results, while smaller windows showed unacceptable instability. We therefore conducted all further tests with 15-bin (30 m) diameter windows, containing a maximum of 177 bins. The parameter under consideration is calculated for the complete window, and the value assigned to the central bin of the window. The window is then moved onwards by one bin, until the end of that line is reached, before being applied to the next row of bins. Since the data are rather gappy in character, we must accept windows which contain less than the full complement of 177 bins. A lower threshold of 50 bins was set, chosen to allow relatively contiguous results from classification, while retaining sensible statistical properties of the window.

An important step in applying surface roughness parameters is firstly to de-trend the data. In the case of a two-dimensional window, this means fitting and subsequently removing a plane to the window, so that the mean of the residuals is zero.

A bewildering variety of surface roughness parameters are defined in the literature, many of which are not useful in the context of ice draft data, particularly a noisy or broken-up dataset as studied here. One useful parameter is the **variogram**; a cornerstone of geostatistics. For any set of data we can compute the variances for every pair of points and then order these according to the distance

(*lag*) between them. The mean value at each lag forms the experimental variogram. The variogram is superior to spectral parameters, since it always exists and is not perturbed by disturbances (data glitches) over small scales. Variograms are typically constituted with 100-200 data points, which matches our chosen 177-bin window.

In geographical applications, the variogram is usually found to have characteristic minima and maxima, and many parameters based on the lag of these maxima and minima are defined. For application to the sonar data, we have to compromise between a full description of the patch (*i.e*. form the variogram over a large area) and spatial resolution in our classification scheme (*i.e*. form the variogram over a small area). The chosen compromise – a 15 bin diameter patch – does not allow the variogram to develop minima – the curve rises smoothly to a 'sill' value which increases gradually thereafter. This is true out to patch diameters of more than 60 bins (120 m), which is far greater than we would choose to classify the swaths. Thus we are limited to more basic properties of the variogram; its maximum value (*pond*) and the slope of the variogram at its minimum lag (*deriv*).

We also develop the second order variogram, (the variogram of the variogram), which is cited as useful if the first order variogram does not display minima. The second order variogram typically allows the detection of shallow large-scale features in the presence of locally-dominating small-scale features, or the detection of features in the presence of strong sampling artefacts, as is the case with the current data.

5.3.1. Separating deformed and undeformed ice

pond robustly segments deformed and undeformed ice. Deformed regions display *pond1* values from 1-10, while undeformed areas have *pond1* values from 0.001 to 0.1. Second order variograms display similar differences between deformed and undeformed ice, with undeformed regions having a minimum value at around 6 m lag.

The plots for variance *vs. pond*, and *deriv vs. pond* suggest that a thresholding value of *pond* will separate the swath into two classes, which, as we have seen previously, we expect to be deformed and undeformed ice. The threshold value in these relations suggests a method of automatically determining an adaptive threshold for the classification by *pond*. We accordingly use the three relations – *pond2*/*deriv2*, *pond1*/variance, *pond2*/variance – to calculate the value of *pond* to split the swath data into deformed and undeformed ice.

5.3.2. Detailed analysis of variograms for regions

Continuing our search for classifiers between FY level and MY level ice, we next examine the character of the variograms for the complete population of undeformed ice bins within any swath. We extract the draft data from swaths containing significant level ice regions, and plot the spread of variogram values at each lag distance. FY level ice has a lower spread of values at each lag, while MY level ice displays the highest values at each lag. This is true for both first order (LH plot) and second order variograms (RH plot). Clearly, the majority of values for both FY and MY ice occupy an overlapping set of values at each lag, however, making it difficult to separate the ice types on a window-by-window basis: the majority of windows containing any given distance/value pair could be drawn from either population.

Extending this analysis to examine a contour plot of the occurrence of any given distance/value pair, ice types may be differentiated by examining the position of the core region with lag: the population of MY ice lag/value pairs has a maximum (dark red area) at very high lags (*c*.15), while the FY core region is seen at lower lag values. This property is maintained across the whole dataset. First order

variograms appear to give better discrimination than the second-order variograms in this respect. This 'high lag population core' criterion holds for all swaths in the current dataset and we therefore propose it as a practical threshold criterion:

> *For first order variograms, a contour plot of the population of lag/value pairs will show a maximum at large lags for MY ice and small lags for FY level ice.*

To explain what is driving this behaviour, the distance at which the maximum population density occurs is the lag at which the family of variograms within the patch all have similar values – *i.e*. the distance/val curves cross over or approach each other more closely. Physically what this means is that MY level ice has a dominant scale of draft variation at dimensions corresponding to the size of our window (30 m), while FY level ice variations occur at smaller scales (2 m for the majority of swaths). This is understandable if we attribute the roughness of MY ice to meltpond relief, the fact that during the summer melt period the melt of the upper ice surface under a melt pond leads to a greater heat flow from the ice bottom and thus a greater bottom melt rate under the same melt pond, causing a pattern of bottom depressions mirroring the pattern of top-surface melt ponds. The lateral scale of these depressions is that of typical melt ponds, i.e. about 30 m, and the undulating ice bottom that is produced is retained through subsequent winters and summers. Meltpond relief indeed produces relatively large scale features. Undeformed FY ice is so flat on the underside that the dominant source of variation in this dataset is probably sonar variability, which is expressed on a bin-to-bin scale (2m in this dataset).

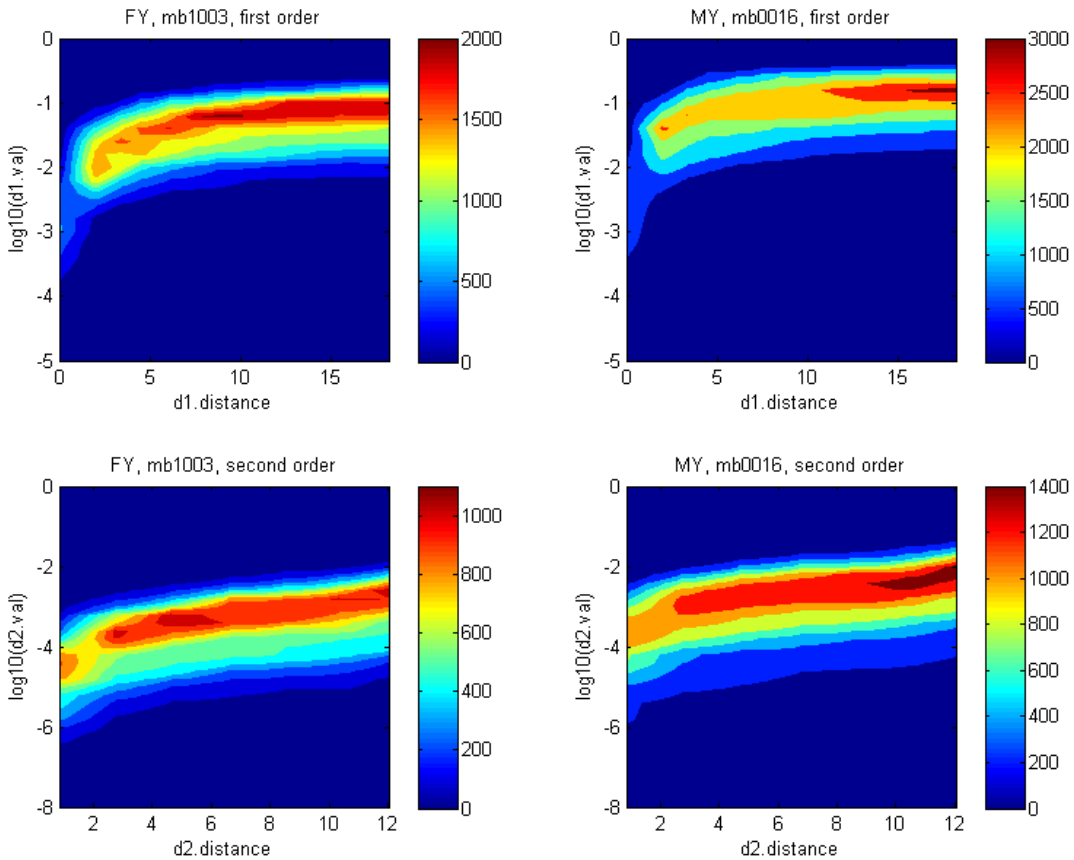

Figure 4.18: Contour plots of the population of data points for each lag/value pair. Results are shown for FY level ice (LH panels) and MY level ice (RH panels). The top two panels are for the first-order variograms, while the bottom two panels are the second-order variograms.

Though not useful for a single window (which only has one value at each distance), this characteristic allows areas of level ice (patches) to be discriminated. We first define regions of level ice using the *pond*/*deriv* criteria then apply the erosion routine to group these into contiguous patches. Each level ice patch can then be separated into FY and MY level ice by examining the lag distance at which its maximum population occurs in lag/value space.

Further confidence in the classification can be gained from the histograms obtained from the distribution of the population by value (*i.e*. collapsing the contour plot of **Figure 4.18** onto the *y*-axis by summing over all *x* distances at each *y* value). MY ice should have a wider range of variogram values at, say, 1/e of the maximum peak height (one-sigma width).

We refine these criteria by running the analysis on all swaths, plotting a single value for the "lag-at-population-max" and "half peak value histogram width" for each swath. Results are promising, with Beaufort Sea and Fram Strait regions generally plotting to lower lags than the North Greenland region. A lag value of around 6.5 appears to be a reasonable choice to segment FY level ice from MY level ice. A similarly 'almost good' division is seen in the width of the global population histograms and we might choose a width of *c*.0.25 as a threshold: all FY swaths lie below this line, though many MY swaths also have widths less than this threshold.

In this two-parameter space, almost all the expected FY swaths fall into the lower LH quadrant – *i.e*. lie under both threshold values. Expected MY ice swaths lie mostly to the right of the dividing line. Lines 1003, 1007, 2013 and 0018 are the only four that appear to be mis-classified by this two-parameter approach. Of course, a single swath can contain different types of level ice, and to refine the classification it is necessary to examine each discrete patch of level ice independently.

5.3.3. Classification by patches

To untangle the effect of mixed ice type swaths, we must perform the classification for each discrete area of level ice. We thus contour the undeformed ice regions and assign a number to each discrete patch within the swath, imposing a minimum patch size (in terms of number of bins within the patch = 100) to allow a robust classification. For each bin in each such patch, we then recalculate the variogram using only the undeformed bins in each window. 'Undeformed' in this respect uses the classification prior to the erosion process, since we seek 'pure' undeformed ice to optimise the classification. The resulting classification is subsequently applied to the larger, eroded, area. The result, for all patches in all swaths is shown in **Figure 4.21**:

The increased level of detail provided in this patch-resolving view indicates that virtually all possible maximum lags exist in each region, though the half power peak width still robustly separates the expected FY and MY ice regions (Fram Strait and Beaufort Sea patches all lie dominantly below the width threshold). The half-power peak width must therefore be considered the dominant level ice classification criterion, with maximum lag only contributing some equivocation in the FY level ice classification. **Figure 4.21** thus suggests that level ice areas can be separated into three types:

1. Peak width > threshold (top half of plot): **MY level ice**
2. Peak width < threshold and lag < threshold (bottom left quadrant): **FY level ice**
3. Peak width < threshold, lag > threshold (bottom right quadrant): mixed age level ice, or data quality issues?

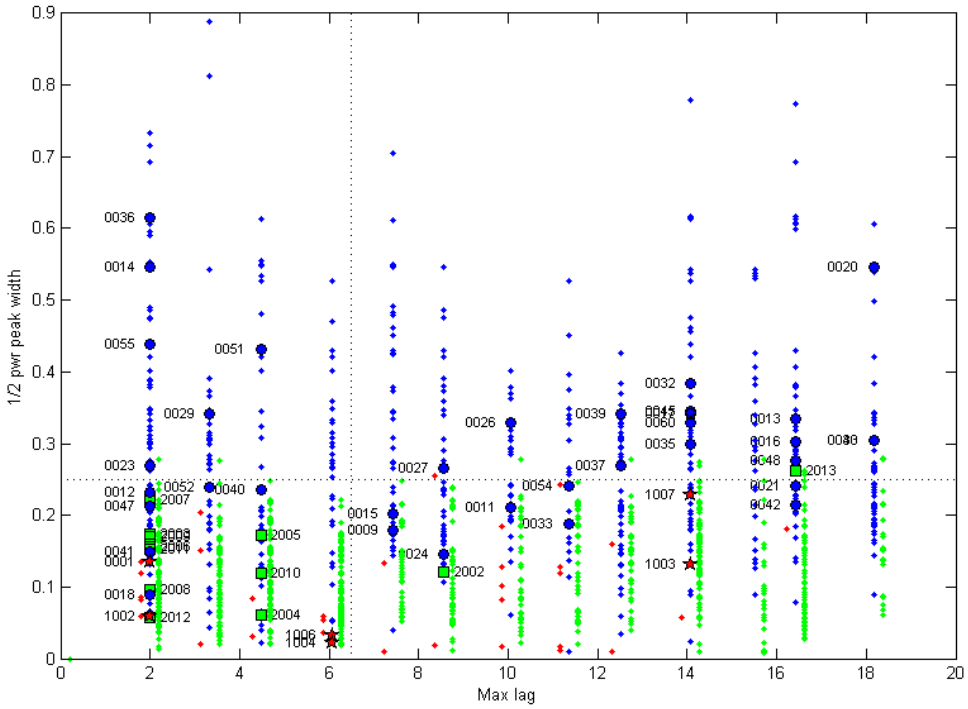

Figure 4.21: Scatter plot of max lag against half power peak width, for all patches in all swaths (dots). Blue dots are for the North Greenland region, green for Fram Strait and red for the Beaufort Sea. Solid symbols are for the whole undeformed ice area in each swath, as in the previous figure. Lags in the three regions are slightly offset for clarity.

To examine the character of the ice in each quadrant, we classify the level ice patches into these quadrants of the **Figure 4.21:**

- Class 1: unequivocal FY level ice (bottom LH quadrant)
- Class 2: probable FY level ice (bottom RH quadrant)
- Class 3: probably MY level ice (top LH quadrant)
- Class 4: unequivocal MY level ice (top RH quadrant)

We also define Class 0 to denote patches which are too small (<100 bins) to generate reliable statistics for classification. To evaluate the ice in each quadrant, we compare this classification scheme (now with seven classes, including deformed and ridged ice) with a three-dimensional view of the interpolated drafts.

5.4. Results

The final classification scheme provides six ice classes:

- Class 0: Level ice, but of a patch size too small to classify reliably
- Class 1: Unequivocally FY level ice
- Class 2: FY level ice, displaying some tendency to MY character
- Class 3: MY level ice
- Class 4: Deformed ice (rubble)
- Class 5: Ridged ice

Classification statistics are summarised in **Table 5.2**, with figures also summed by region. *Beaufort Sea* data show dominantly FY level ice (24% + 21%), with relatively little level MY ice (6%). Deformed ice (rubble + ridge) makes up 44% of the area – a typical figure for the region (*e.g*. Melling *et al*., 1995) and identical with that found in the same area at the same time using a small *Gavia* AUV (Doble *et al*., 2011). Ridging makes up 15% of the total area.

The *North Greenland* region shows slightly more MY level ice (26%) than FY level ice (22%). Deformed ice dominates the region (class 4+5 = 48%), as expected.

Results for the *Fram Strait* region suggest that almost no MY level ice is present (1%), with FY level ice dominating (67%) and relatively little deformed ice (26%). We note that this region has the poorest data quality, however, as the submarine is running relatively fast (covering 15 km track length in 3800 pings – other regions typically populate 4-8 km track length with the same number of pings) and deep (143 m). The classification is thus particularly challenging and suggest that these data are simply too sparse to be reliably classified.

Figure 5.X a) Example classified swaths from the Beaufort Sea (mb1003 top, mb1007 middle)

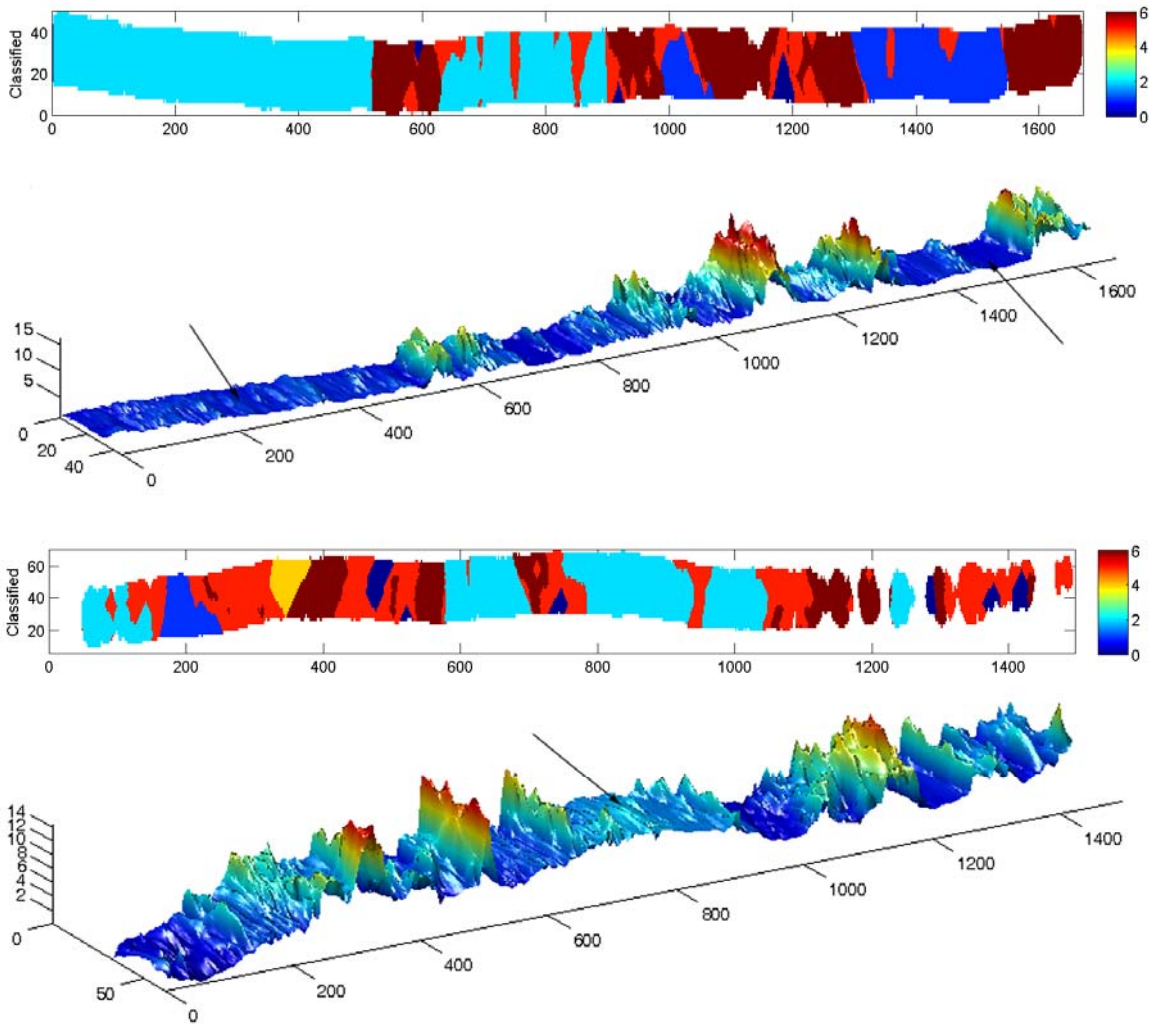

Figure 5.X c) Example classified swath from the Fram Strait (mb2013)

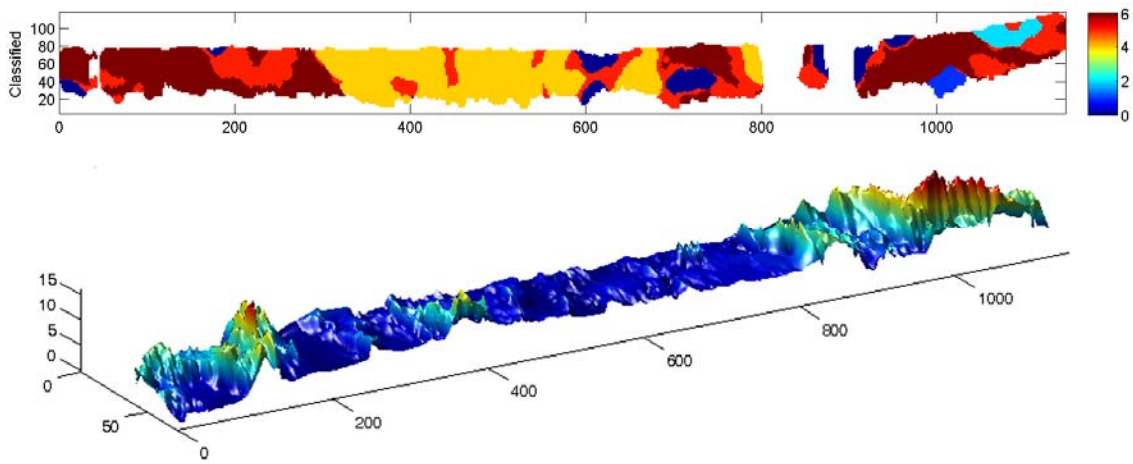

Table 5.2: Summary statistics (percent) for each swath by ice class (0-5). Asterisks next to the swath number indicate that the alternative, change-in-slope, method was used to determine the P2D2THRESH value for that swath (4 cases).

		Unc level	FY level	FY' level	MY level	Def	Ridged
Swath	**Total area, m²**	**0**	**1**	**2**	**3**	**4**	**5**
Beaufort Sea							
mb0001	34,028	7	0	0	40	34	13
mb1002	45,440	1	76	0	0	12	10
mb1003	56,144	1	16	45	0	10	28
mb1004	51,712	6	22	14	0	54	4
mb1006	8,520	8	23	15	0	33	21
mb1007	47,964	4	5	36	3	34	18
Region	**243,808**	**4**	**24**	**21**	**6**	**29**	**15**
North Greenland							
mb0009	43,524	11	9	13	0	46	20
mb0011	27,380	2	0	5	0	24	69
mb0012*	91,764	6	19	3	0	46	25
mb0013	251,816	5	2	0	39	36	18
mb0014	367,908	3	0	1	63	28	5
mb0015	116,288	5	7	33	0	32	23
mb0016	189,284	4	0	4	36	38	18
mb0017	280,940	3	7	11	33	29	16
mb0018	39,380	13	26	0	0	36	26
mb0020	341,828	4	3	1	57	31	5
mb0021	128,028	3	19	35	8	24	11
mb0023	343,996	4	0	12	24	40	20
mb0024	114,372	5	6	38	0	34	18
mb0026	326,948	6	0	0	31	39	25
mb0027	111,664	3	5	10	41	28	12
mb0029*	336,464	3	9	2	36	37	13
mb0030	99,564	4	2	23	7	56	9
mb0032	331,408	5	11	2	30	43	9
mb0033	167,072	3	13	38	0	33	12
mb0035	343,152	6	11	15	12	44	12

mb0036	172,548	3	10	6	49	31	1
mb0037	206,436	5	16	20	1	45	13
mb0038	40,440	3	0	0	47	33	16
mb0039	225,336	5	5	10	27	40	13
mb0040	240,720	5	11	15	4	39	27
mb0041	62,956	3	19	27	0	36	16
mb0042	387,748	3	4	37	13	32	11
mb0043	140,216	3	0	24	37	23	12
mb0045	211,432	4	1	21	36	27	11
mb0047	17,032	1	0	30	0	27	42
mb0048	56,776	7	9	15	8	48	13
mb0049	33,684	1	22	35	0	17	26
mb0051	120,536	4	20	0	45	29	3
mb0052	197,060	2	15	42	6	25	10
mb0054	319,064	4	10	28	16	18	24
mb0055	110,296	3	4	27	33	21	12
mb0060	37,116	8	15	0	8	49	20
Region	**6,632,176**	**4**	**7**	**15**	**26**	**34**	**14**

Fram Strait

mb2002	409,744	8	10	42	0	29	10
mb2003	130,720	5	18	13	0	30	34
mb2004	65,548	7	24	23	0	32	13
mb2005*	478,564	6	25	34	1	29	5
mb2006	523,820	5	50	26	0	18	1
mb2007	523,748	3	42	36	6	9	3
mb2008	504,456	6	36	24	0	19	14
mb2009*	611,492	4	43	37	0	4	13
mb2010	599,092	4	36	40	0	17	3
mb2011	518,120	5	51	22	0	15	6
mb2012	245,048	5	35	31	0	27	2
mb2013	57,416	7	2	3	30	23	35
Region	**4,667,768**	**5**	**36**	**31**	**1**	**18**	**8**

5.5. Discussion

The 2007 voyage of *Tireless* was a unique (to date) opportunity to generate multibeam 3D sonar swaths of the ice underside over a complete transect of the Arctic Ocean, with sufficient variety of ice types to facilitate the development of an ice classification system for multibeam, and sufficient size of the dataset to require such a system for interpretation of ice characteristics. The quality of the data, however, for reasons already explained, was not as high as was hoped for.

Though the absolute draft data cannot be relied upon, the (uncompensated) submarine motion occurs on length scales far greater than the scale of our surface roughness analysis. It has, at least, forced us to develop a system that relies totally on these roughness characteristics, without any support from ice draft information. The lack of reliable depths in fact only affected our ability to separate open water/refrozen leads from flat FY level ice. The success of the classification scheme with this relatively-poor quality data augurs well for its application to properly calibrated and processed surveys, where we expect that the task will be significantly simpler.

Data were obtained from three very different regions, representing the complete range of ice conditions liable to be encountered: from the thin, relatively undeformed ice cover of the Beaufort Sea, through the formidably deformed MY ice regime north of Greenland, into the Fram Strait outflow region, where ice of all types can be expected.

We were able to demonstrate reliable segmentation of undeformed, deformed (rubble) and very deformed (ridged) ice, using the variogram properties (magnitude, slope) and variance within a single window. Separating FY level ice from MY level ice required rather more application, since their variogram characters are sufficiently similar that considerable overlap of the magnitude values exists between the two ice types. This precludes segmentation using the properties within a single window and we must instead group the variogram properties together inside a contiguous area of undeformed ice. The 'family' of variograms within a patch of undeformed ice (*i.e*. a series of curves relating variogram magnitude to separation distance, or lag) have two characteristic properties which allow FY and MY level ice to be separated. We first form a contour plot of the population density of values in each value/lag bin. The maximum population density (the most populated bin) tends to occur at low lags for FY ice and higher lags for MY ice. The determining threshold is lag = 6.5 for this dataset. We then collapse the contour plot onto the value axis, which forms a probability density function (PDF) of population against variogram value. The width of the population peak, specifically the one-sigma width, is then used to separate ice types: FY level ice has a much narrower spread of values than MY ice, with the dividing threshold being around 0.25 for these data.

Both criteria find ready explanation in the physical properties of the ice cover: MY ice is both more variable in draft (hence the wider one-sigma width) and has a characteristic scale of draft variation (due to meltpond relief) greater than for FY ice – hence the higher value of lag at which the population concentrates.

The final result is that in the two reliable data regions studied – north of Greenland and the Beaufort Sea, the areal percentage of deformed ice is very high – around 44-48% - and almost exactly matches the small-scale study of Section 4, as well as earlier surveys in the literature. This result shows that CryoSat's demonstrated underestimation of ice thickness in deformed ice regions is a serious problem, even in the Beaufort Sea, and will lead to the sensor underestimating the overall volume of this sea ice type in the Arctic.

6. Discussion & Conclusions

The recent availability of the AWI CryoSat Level-2 data has greatly improved on the quality of altimeter data available to researchers who do not wish to become involved in the minutae of retracking individual radar waveforms. This dataset also preserves the various assumptions made about snow properties, allowing easy manipulation of the results. The assumptions are reasonable and informed by the investigators' unrivalled experience in performing *in situ* validation campaigns with co-incident airborne radar and EM instruments, coupled with the latest findings on snow depth over sea ice, from the extensive NASA Ice Bridge flights.

Despite all this, when the actual ice thickness from the CryoSat radar is compared with large-scale measurements (**Section 3**), we find that the inaccuracy greatly exceeds the "sea ice thickness uncertainty" parameter embedded in the AWI dataset. Comparison with the BGEP moorings should be 'as good as it gets', since both CryoSat and the mooring are dealing with monthly averaged parameters at a specified point, but the error across the 2011-2012-2013 periods is 0.98 m RMS, 0.40 m mean. This is of the same order as that predicted by radiometric modelling of snow thickness and density variability. Over-estimation by CryoSat is perhaps also to be expected conceptually, since any areas of very thin ice are likely to include open-water and/or leads, both of which result in the CryoSat waveform to be rejected from the ensemble analysis.

Comparison with Ice Bridge tracks was less revealing, showing a general over-estimation of CryoSat ice thickness, but no clear pattern with latitude, data quality or ice type. Previous investigators have found the comparison equally baffling. We suggest that this is a combination of (a) because neither technique can tell us the actual ice thickness, since both are measuring a very small number (the freeboard) and converting this to thickness using the assumptions detailed in Section 1; (b) the fact that very small areas of flat ice (in the range 0.2 – 16% of area) can dominate the backscatter return (Tonboe *et al*., 2010 – see Figure 8 there) – termed *preferential sampling*.

Similarly, comparison with HEM results is only reliable over large regions of level ice, since the EM thickness over deformed ice is a complex convolution of ice topography, instrument footprint and porosity/seawater inclusions, which is currently poorly constrained. From our analyses, CryoSat appears to under-estimate ice thickness over deformed ice. It should be borne in mind that the EM technique also under-estimates absolute ridge drafts (compounding the suggested CryoSat errors), though the overall effect on a footprint-averaged basis has yet to be assessed. The "flat ice – high backscatter effect" will also play a role here.

Broadly speaking, CryoSat appears to over-estimate the thickness of level ice and under-estimate the thickness of deformed ice. While the mechanism for the under-estimation of level ice can be understood, the returns over deformed ice require more investigation.

Towards this end, the full 3D co-incident top- and bottom-surface study in **Section 4** is instructive. We find that the draft-freeboard ratio increases markedly over deformed ice features, ranging from 1 over thin, snow-covered refrozen leads to almost 10 in the FY ridge feature. The high *R* values over deformed features stem from the fact that these features are much wider underwater than their surface expression, plus the fact that the submerged portion of the deformed ice is porous to seawater, increasing the overall density of the feature if the overall envelope of the keel is considered. The *in situ* data also demonstrate the significant departures from climatological snow properties which occur at CryoSat footprint scales: the AWI reprocessor suggests 16 cm snow

thickness and 320 kg/m^3 snow density for this region, while the measured properties over the 300x300m areas were 30 cm and 240 kg/m^3 respectively.

The draft/freeboard relation appears to be rather stable with increasing footprint size, from patches of 11m diameter to the full dataset, which should give some reassurance of the representativeness of large footprint measuring techniques.

Section 5 establishes the ubiquity of these mixed-ice type footprints, with an almost 50/50 split between deformed and level ice in the Beaufort Sea and North Greenland regions. Additionally, very few swaths contained one class of level ice, indicating surface roughness variations (and hence preferential sampling issues) even in the level ice regions.

The main stumbling block to progress, is the response of the CryoSat radar to the mixed ice types and deformed ice thicknesses in its footprint – "preferential sampling". Radiometric modelling suggests that this response is dependent on surface roughness, with the high backscatter magnitude from the thinnest ice within the footprint largely determining the elevation of the effective scattering surface.
For instance *Tonboe et al* (2010) modelled the radar response to mixed ice types in a series of CryoSat footprints taken from co-incident Radarsat scenes and EM thickness transects. They demonstrated that a real ice thickness mode of 2.4 m (from the EM) would be rendered as 1.0 m in a Fram Strait scene. This effect is at least as important as the better-known snow property errors, but is not discussed in any subsequent literature.

The problem is that - while we can forward-model from a comprehensive overview of ice type and thickness, as done by Tonboe *et al*., to determine the response of CryoSat - it is not possible to backward-model a given CryoSat ice thickness value and generate the correct ice thickness distribution. CryoSat appears to be limited in usefulness to those of footprints which sample a single ice type. Clearly, in most regions, this type of footprint is rare.

We are in the process of preparing a publication from these findings, and will further examine the detailed tracks of full-resolution mooring, EM and Ice Bridge data, to better understand the scattered correlation between CryoSat and these measurements in terms of preferential sampling.

7. References

Ackley, S. F., W. D. Hibler III, F. K. Kugzruk, A. C. Kovacs, and W. F. Weeks (1976), Thickness and roughness variations of Arctic multiyear sea ice*Rep. 76-18*, 25 pp, CRREL, Hanover.

Doble, M. J., P. Wadhams, A. Forrest, and B. E. Laval (2009), Through-ice AUV deployment: Operational and technical experience from two seasons of Arctic fieldwork, *Cold Reg. Sci. & Tech.*, *56*, 90-97.

Forsberg, R., and S. Kenyon (2004), Gravity and geoid in the Arctic region - the northern gap now filled, paper presented at GOCE workshop, ESA-ESRIN.

Forsberg, R., and H. Skourup (2005), Sea ice thickness from airborne laser altimetry over the Arctic Ocean north of Greenland, *Geophys. Res. Lett.*, *29*(20), 1952-1955.

Forsstrom, S., S. Gerland, and C. A. Pedersen (2011), Sea ice and snow densities and hydrostatic equilibrium assumption from in situ measurements in Fram Strait, Barents Sea and Svalbard coast, *Ann. Glaciol.*, *57*, 216-270.

Haas, C., S. Hendriks, and M. J. Doble (2006), Comparison of the sea ice thickness distribution in the Lincoln Sea and adjacent Arctic Ocean in 2004 and 2005, *Ann. Glaciol.*, *44*, 247-252.

Hvidegaard, S. M., and R. Forsberg (2002), Sea-ice thickness from airborne laser altimetry over the Arctic Ocean north of Greenland, *Geophys. Res. Lett.*, *29*(1952), doi: 10.1029/2001GL014474.

Hvidegaard, S. M., R. Forsberg, and H. Skourup (2006), Sea ice thickness estimates from airborne laser scanning, in *Sea Ice Thickness: Past, Present and Future*, edited by P. Wadhams and G. Amanatidis, pp. 193-206, Brussels.

Kovacs, A. C. (1997), Estimating the full-scale flexural and compressive strength of first-year sea ice, *J. Geophys. Res.*, *102*(C4), 8681-8689.

Krabill, W. B., R. H. Thomas, C. F. Martin, R. N. Swift, and E. B. Frederick (1995), Accuracy of airborne laser altimetry over the Greenland ice sheet, *Int. J. Rem. Sens.*, *16*(7), 1211-1222.

Kurtz, N. T., and S. L. Farrell (2011), Large-scale surveys of snow depth on Arctic sea ice from Operation IceBridge, Geophys.Res. Lett., 38, L20505, doi:10.1029/2011GL049216.

Kurtz, N. T., T. M. Markus, D. J. Cavalieri, W. B. Krabill, J. G. Sonntag, and J. Miller (2008), Comparison of ICESat lidar data with airborne laser altimeter measurements over Arctic sea ice, *IEEE Trans. on Geosci. and Rem. Sens.*, *46*(7), 1913-1924.

Kwok, R., G. F. Cunningham, H. J. Zwally, and D. Yi (2007), Ice, Cloud and land Elevation Satellite (ICESat) over Arctic sea ice: Retrieval of freeboard, *J. Geophys. Res.*, *112*(C12013), doi: 10.1029/2006JC003978.

Laxon S.W., K. A. Giles, A. L. Ridout, D. J. Wingham, R. Willatt, R.Cullen, R. Kwok, A. Schweiger, J. Zhang, C. Haas, S. Hendricks,R. Krishfield, N. Kurtz, S. Farrell and M. Davidson (2013). CryoSat-2 estimates of Arctic sea ice thickness and volume, *Geophys. Res. Lett*., 40, 732–737, doi:10.1002/grl.50193.

Lepparanta, M., and R. Hakala (1992), The structure and strength of first-year ice ridges in the Baltic Sea, *Cold Reg. Sci. & Tech.*, *20*, 295-311.

Melling, H., D. R. Topham, and D. Reidel (1993), Topography of the upper and lower surfaces of 10 hectares of deformed sea ice, *Cold Reg. Sci. & Tech.*, *21*, 349-369.

Tonboe, R. T., L. T. Pedersen, and C. Haas (2009), Simulation of the satellite radar altimeter sea ice thickness retrieval uncertainty, *The Cryosphere Discuss.*, *3*, 513-559.

Tonboe, R. T., L. T. Pedersen, and C. Haas (2010), Simulation of the Cryosat-2 satellite radar altimeter sea ice thickness retrieval uncertainty, *Can. J. Rem. Sens.*, *36*(1), 55-67.

Wadhams, P., and M. J. Doble (2008), Digital terrain mapping of the underside of sea ice from a small AUV, *Geophys. Res. Lett.*, *35*(L01501), doi: 10.1029/2007GL031921.

Wadhams, P., W. B. I. Tucker, W. D. Krabill, R. N. Swift, J. C. Comiso, and N. R. Davis (1992), Relationship between sea ice freeboard and draft in the Arctic Basin, and implications for ice thickness monitoring, *J. Geophys. Res.*, *97*(C12), 20,325 - 320,334.

Warren, S. G., I. G. Rigor, N. Untersteiner, V. F. Radionov, N. N. Bryazgin, Y. I. Aleksandrov, and R. Colony (1999), Snow depth on Arctic sea ice, *J. Clim.*, *12*(6), 1814-1829.

Wingham, D. J., et al. (2006), CryoSat: a mission to determine the fluctuations in the Earth's land and marine ice fields, *Adv. in Space Res.*, *37*, 841-871.

Giles, K.; Laxon, S.; Wingham, D.; Wallis, D.; Krabill, W.; Leuschen, C.; McAdoo, D.; Manizade, S. & Raney, R. Combined airborne laser and radar altimeter measurements over the Fram Strait in May 2002 Remote Sensing of Enviroment, 2007, 111, 182-194