We would like to thank the anonymous reviewer for the very useful comments and suggestions which help us improve the quality of our paper. Before replying to the more specific points raised by the reviewer we wish to clarify a few general points.

The main goal of our study is to provide new estimates of uncertainties in Arctic sea ice thickness and volume derived from space. Our assessments of the uncertainties are resulting from uncertainties in geophysical parameters such as sea ice area, snow depth and sea ice density. As stated by the reviewer the main challenge is to understand how well errors are correlated in space, time and with each other and to what degree they are random. In our paper we present how these correlations can be considered in the uncertainty assessment on large scales, taking into account how the geophysical parameters are retrieved and used for calculating sea ice thickness and volume.

The reviewer acknowledges that the overall goal of our paper is good, but requests more information about the methodology used. The main concern raised by the reviewer is a lack of a deep description of the above mentioned correlations and how they are really considered with our Monte-Carlo approach. The reviewer is indeed right and the method has not been described in sufficient detail. We have addressed the raised issues in the following sections and will include the missing information in our paper.

In many previous studies (e.g Giles et al. 2008) uncertainties in sea ice thickness are calculated with a classic error propagation with the following formula:

$$\underbrace{\sigma_{total}^2}_{\text{total uncertainty of sea ice thickness}} = \underbrace{\sigma_{fs}^2\left(\frac{\rho_w}{\rho_w-\rho_i}\right)^2}_{\text{freeboard}} + \underbrace{\sigma_{h_s}^2\left(\frac{\rho_s-\rho_w}{\rho_w-\rho_i}\right)^2}_{\text{snow depth}} + \underbrace{\sigma_{\rho_i}^2\left(\frac{h_s(\rho_s-\rho_w)+f_s\rho_w}{(\rho_w-\rho_i)^2}\right)^2}_{\text{ice density}} + \underbrace{\sigma_{\rho_s}^2\left(\frac{h_s}{\rho_w-\rho_i}\right)^2}_{\text{snow density}}$$

where $h_s$ is the snow depth, $f_s$ the (snow-ice) freebord, $\rho_w$, $\rho_i$, $\rho_s$ the densities of water, ice and snow, and $\sigma_l$, $\sigma_{fl}$,$\sigma_{hs}$, $\sigma_{\rho i}$, $\sigma_{\rho s}$ are the uncertainties (standard deviations) of the sea ice thickness, freeboard, snow depth, density of ice and snow. Hereby uncertainties are treated as uncorrelated and values are taken from previous in-situ measurements as described in literature. By uncorrelated we here mean that there is neither a correlation between the parameters considered nor spatial auto-correlations.

In our approach the correlation between the parameters has been included by varying all parameters simultaneously in our Monte-Carlo run. These resulting total uncertainties of sea ice thickness and volume are given in Figure 6, 8 and 9. The relative contribution (shown in different colors in figure 6 and 9) resulting from the three parameters are calculated as a relative contribution of the separate sensitivity calculations to the total uncertainty.

To include a spatial auto-correlation in our method we changed the values 'globally', so we used one random value from the probability distributions for all pixels of first-year-ice and another one for multi-year-ice (or weighed by the fraction of multi-year ice where needed). For snow we additionally included a connection between first-year-ice and multi-year-ice by using half of the value picked for multi-year-ice for first-year-ice. With this method the auto-correlation is given only by the dependency on the sea ice type, which is clearly a simplification of the true nature. It is, however, a first step in the right direction considering how this calculation is performed. In our study (as in many others) the snow depth has been taken from a climatology, or to be precise a modification of a climatology. The primary source of uncertainty is therefore a bias resulting from an over or underestimation of the snow depth in general while regional uncertainties or sub-gridscale variability (as analysed e.g. in Kurtz et al 2009) are of second order importance. Our analysis would however look different if snow depth had been estimated from accumulation or passive microwave retrievals. For the density of ice there is no consistency in the mean value used in the most cited publication (see table 3 for examples). As pointed out in our manuscript this has substantial influence on assumptions on long-term trends in sea ice thickness and volume. A random natural variability of the values for sea ice density should be included in the future. But, considering the lack of knowledge of our community in that particular point, this random natural variability is of second order importance.

---

Below we answer on the more specific questions and comments from the reviewer:

*The paper also could be helped by more clearly defining what is meant by the uncertainties. The authors state that the uncertainties are for "mean ice thickness" or "total volume" but don't state over what time averaging periods. Are those for individual IceSat campaigns or aggregate over multiple years.*
→ All uncertainties should be understood as uncertainty per IceSat campaign, we will clarify this in the paper.

*I wonder if the paper could be restructured so a "best guess" estimate could be presented first, which is then followed by the uncertainty assessment of this specific first guess. I think that might help with reading this.*
→ We followed this approach for the sea ice volume, while for the sea ice thickness this is more manifold. We show both the range of values for the 'absolute' sea ice thickness (Figure 3) and the uncertainties from the Monte-Carlo approach for the effective sea ice thickness (Figure 6). For the range of values for sea ice thickness we provide a kind of 'buest guess' but only based on the methods used in literature (brown line in figure 3 and 5). This is slightly different from our Monte-Carlo approach, that we believe is better, so we didn't label it as best guess to avoid confusion. We will add some information about this in the revised version.

*The paper suggests that the Laxon 2013 findings about the ice volume loss between the IceSat period and the CryoSat period are substantially affected by the different assumptions about ice density. Since IceSat draft retrievals have been directly compared with in-situ measurements, the question arises how those comparisons would look with different density assumptions.*
→ The reviewer is touching a very important point by raising this question. As shown in our paper a correct and consistent choice of the density is essential to be able to draw any conclusions for comparing different datasets and therefore to validate them. Kwok at al. 2008 concluded that the IceSat estimated drafts seem to be slightly overestimated in fall and underestimated in winter. In fall

the lower density would reduce the difference, while in winter it would become larger. In winter however, the snow load is high and may influence the results. In general the use of ULS measurements for validation of space borne data may be questioned: On one hand ULS drafts tend to overestimate the mean (or modal) sea ice thickness as reflection from ridges results in a positive bias. Space borne altimeters, on the other, are not able to resolve the ridges and therefore rather underestimate the mean ice thickness.

*Are there other substantial differences in the Kwok et al. 2009 and Xi an Zwally data sets that need to be considered?*
→ The substantial difference between the two datasets is that Kwok et al. 2009 included corrections for the reflectivity used to identify tie-points (leads). Corrections have been included for two reasons: 1) leads with high reflectivity might be covered by layers of snow that biases the elevation estimates 2) leads are often not covering the whole IceSat footprint, and the surrounding ice may again bias the elevation estimate.
We were however not able to compare the freeboard retrievals directly, but compared the thickness estimates using the same density of sea ice in figure 3. Using the same densities of 925 kg/m³ the results are relatively close (black line, D2). The corrections included in the JPL dataset increase the freeboard while the snow accumulation product used by Kwok et al 2009 seems to give a higher snow depth than in the modified Warren climatology (see figure Kwok 2009, Figure 5 b ). This lowers the thickness estimates. For the total mean the two effects therefore cancel each other to some extent, but as none of the datasets is available we are not able to make quantitative conclusions.

---

*5053 lack of sufficient estimate of uncertainties.. Please clarify what is meant? What would suffice?*
→ A sufficient error analysis would include the above mentioned correlation in space and time and the correlation between the parameters. At the same time the way sea ice thickness is calculated has to be considered. The use of standard deviations from field campaigns as random error for calculating uncertainties in sea ice thickness and volume (as specified above and done e.g. in Giles et al 2007, Alexandrov et al 2010, Kwok et al 2009) does not include any of these points. We will rewrite and clarify this point raised by the reviewer.

*5043, spatial autocorrelation This is mentioned here it is a very good point but it is never really picked up again as far as I can tell.*
→ See discussion above. We will revise section 3.4 and other parts accordingly were appropriate.

*5055.. for comparison we also use the gridded data set from JPL I don't see this comparison anywhere except in maybe Fig 8. Are those the JPL thickness retrievals? If so, how is the mean volume calculated there. Is this "weighted" by ice concentration or not.*
→ An important comparison is also given in Figure 3. Using the same assumptions for sea ice density as used in the JPL the values closely resemble each other (see discussion above). For the values shown in Figure 8 the sea ice concentration is considered.

*5057, 3, frequency spectrum This isn't strictly correct since many sea ice algorithm stake advantage of the polarization differences so maybe more generically "frequency and polarization" signatures.*
→ This is indeed correct. We will modify the text accordingly.

*5061 explain more clearly how this is done and specifically how this addresses the shortcomings of other error analyses (e.g. spatial auto-correlation of errors, correlation of errors in time, correlation of errors between parameters)*
→ This point is discussed above.

*5064, 26 sub-grid scale variability How is this computed? What does this mean in this context.*
→ In figure 5 we calculate the sea ice thickness in 5 different ways. Two of them have an ice type dependent snow depth (S2 and S3): First (S2) we assumed the pixels to be either FYI or MYI. For MYI the snow depth was chosen according to the Warren climatology, and was reduced by 50% over FYI. For the second approach (S3) we weighted the snow depth from the Warren climatology by the MYI fraction in each pixel. The resulting mean sea ice thickness in the Arctic calculated with these two methods differs by a few cm only. We will rephrase the sentence to clarify this point.

*5066,5 no increase in winter ice thickness This is interesting. Kwok and Rothrock found that the IceSat estimates substantially underestimate the seasonal cycle. Could this be part of the reason? Kwok uses model predicted accumulation? How different is that relative to the reweighted W99 climatology.*
→ It is indeed a rather surprising result. We believe that this is mostly driven by the overestimation of snow depth in winter and actually consistent with the finding by Kwok and Rothrock. The mean snow depth from the simulated snow accumulation used by Kwok, however, is slightly larger in spring (see Kwok et al 2009, Figure 5b ), but also varying from year to year. Additionally the spatial distribution is different (see Kwok and Cunningham, Figure 4b) so it is hard to make quantitative statements. Also, it seems that there are other factors (such as the above mentioned corrections for the freeboard) which influence the results. However, an evaluation of the spatial and temporal variability of the simulated snow accumulation compared to Warren (1999) should be done to understand this properly. As far we know, this dataset is not available, unfortunately.

*5065, 29 mean absolute uncertainty Clarify over which time period this is, single day, one campaign, multiple campaigns?*
→ The uncertainty is calculated per campaign. We will clarify this in the revised text.

*5072, 28 rate of -385 km3/a.. This is much closer to the PIOMAS-derived volume loss than Kwok et al. or Laxon et al. Though the differences in length of time series should be noted.*
→ The value is indeed closer. As the timescale is different and very short, we believe that a comparison does not add useful, and more important, solid information.

*5074, 16.. the lack of sensitivity to ice concentration errors is somewhat surprising, can this be explained in more detail?*
→ The uncertainties in sea ice area occur mostly in those areas where the sea ice is very thin like in the marginal ice zone and close to land. For the extent or area this has a large influence but it does not

propagate into the estimates of sea ice volume.

*5075, 15 However, What are you saying here? The uncertainties shown in this paper suggest that IceSat or Cryosat-2 derived changes are nonsense? I may agree with that point but you don't make it very well. You show that there is about a 30% error in the volume trends. This is a quantitative conclusion. I think some rewording of the idea is needed.*
→ The comparison between the sea ice volume resulting from ICESat and CryoSat-2 has not been done in a fully consistent way and we believe that in our study we made a more consistent comparison. We accounted for the differences in sea ice density while e.g. the snow depth is not considered. Overestimation of snow depth, and more importantly no consideration of the partial penetration of the radar signal into snow, may result in an overestimation of the volume from CryoSat. We therefore believe that more work is required before making any strong quantitative statement about the long-term changes in sea ice volume. In particular a quantitative statement in the conclusions would be misleading, but for a more interested reader the values are provided in the discussion section.

References:

Alexandrov, V., Sandven, S., Wahlin, J. & Johannessen, O. :The relation between sea ice thickness and freeboard in the Arctic, The Cryosphere, 2010, 4, 641-661

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

Kurtz, N. T., Markus, T., Cavalieri, D. J., Sparling, L. C., Krabill, W. B., Gasiewski, A. J. & Sonntag, J. G.: Estimation of sea ice thickness distributions through the combination of snow depth and satellite laser altimetry data, Journal of Geophysical Research: Oceans (1978--2012), 2009, 114

Kwok, R., Cunningham, G., Wensnahan, M., Rigor, I., Zwally, H. and Yi, D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003—2008, JGR, 2009, 114

Kwok, R. and Cunningham, G. ICESat over Arctic sea ice: Estimation of snow depth and ice thickness JGR, 2008, 113, C08010