Reply to anonymous referee #1

We thank the anonymous referee for many useful comments, suggestions and questions.

General comments.
We agree with the referee that it would be a very interesting study to compare all of the published methods for deriving elevation changes from ICESat data. But it is not the purpose of this manuscript to do a review of published methods, but just to highlight the fact that different methods indeed does produce different results.

We thought it to be interesting exactly how much the cross-over analysis (M4) differs from the along-track methods (M1-M3). But we agree that M4 will never ”win” in a comparison, and we will not include this as a separate method in a revised manuscript. We will instead include the cross-over points in the M1-M3 analysis, which in the current form do not take cross-overs into account but are along-track analysis only. This can easily be done.

We also agree that the editing based on variances is not the best way to judge the performance of the methods, and we would not use this approach in a revised manuscript.

Several question from the referee are related to error propagation.
One very important point which must be clarified in a revised manuscript, is that we do *not* perform a full error propagation through the dH/dt calculation. We assume equal error on the measurements in the individual along-track segments, and estimate the variance from the least squares regression. The bootstrap approach is used to determine the error estimate on the volume change.

Several additional figures and explanations are requested by the referee, but considering the length of the manuscript, we propose that some of the extra material is presented in supplementary material to a revised manuscript.

Specific comments
2. ICESat data

-We propose to add the following sentences to Section 3 (p.2108, l7):
”An observed difference will also include components which are not related to the ice sheet mass balance. The compaction of the firn, vertical bedrock movement caused by glacial isostatic adjustment (GIA), and present-day mass changes all cause elevation changes which is part of the observed elevation difference. Furthermore a potential elevation bias between the laser campaigns must be considered, since this would also be interpreted as elevation changes.”

-The basal melt of the ice sheet actually does contribute to the mass balance, and this issue is discussed in further detail in the reply to anonymous referee #2

2.1 ICESat data pre-processing

- The sentence (p.2107, l24) ”We have also used the difference between the shape of the return signal and a Gaussian fit (the IceSVar parameter), to evaluate data. Large differences indicate less reliable surface elevation estimates..” should be replaced by ”We have also used the standard deviation of the difference between the shape of

the return signal and a Gaussian functional fit (the IceSvar parameter), to evaluate data. Large standard deviations indicate less reliable surface elevation estimates..”
The IceSvar parameter (of 0.035V) was used in the data culling in:
Reference: Smith, B. E., Fricker, H. A., Joughin, I. R. and Tulaczyk, S. : An inventory of active subglacial lakes in Antarctica detected by ICESat (2003-2008), Journal of Glaciology, vol 55, 2009.
We use a threshold of 0.04V in this study.

- We propose to include information in the text, on the percentage of data removed by the individual criteria, instead of adding too many details to table 1.
”In the data culling, 78.4% of the rejected data are rejected by various quality flags, 21.1% by the IceSVar parameter and only 0.5% are rejected by the number of peaks criterion.”

-We will rephrase sentence p. 2108, l. 2-3 to ”Multiple peaks can be caused by reflections from clouds and by topography in the illuminated footprint”.

- We have observed the crossover differences (dH) within the individual laser campaigns and found that unrealistically large dH values were removed after data culling. We propose to provide a more detailed discussion of this issue in the supplementary material of a revised manuscript.

-The surface type definition is adapted from:
http://websrv.cs.umt.edu/isis/index.php/Present_Day_Greenland
Bea Csatho, Toni. Schenk, C.J. van der Veen, William B Krabill, Presented at the AGU 2009 Fall Meeting

3 Methods for deriving surface elevation changes

-We propose to rephrase the sentence (p.2108, l. 17-20) to:
”The fact that the ICESat measurements are not exactly repeated, complicates the methods for deriving dH/dt. Any separation between two measurements introduces a surface slope component, which can be decomposed into an along-track and cross-track component.”

- The 500 m segment size ensures that 2-3 measurements from each campaign is present, which is the minimum needed to solve the governing equations. By choosing this relatively small segment size, the assumption that all measurements in each segment are associated with equal uncertainty is more valid than if the measurements were separated by a larger distance (more on this later). We wanted to get the most out of each of the methods presented, and did not want to degrade M2 and M3 to make the comparison to M1 more fair.

-We agree, and propose not to use the variance to edit the dH/dt estimates in a revised manuscript, and not to judge the method performance based on this criterion.

Some generic comments for methods M1-M4

-The questions about the error propagation are very important.

A major point is that we have not done a full analytical error propagation. We assume that the error of the measurements (and of the DEM used in M1) in each segment is the same, and determine the variance from the least squares regression.
The volume uncertainty is determined by the bootstrap approach, which is an alternative to the full error propagation in situations where this is difficult or even impossible to do analytically.
This will be emphasized in a revised manuscript.
A more detailed description of the bootstrap approach is given later.

-We propose to show figures of the variances in supplementary material to a revised manuscript..
.
-As mentioned earlier, the cross overs points are not included in M1-M3 at the moment, and we propose to exclude M4 but instead include this information in M1-M3.

3.1 Method 1.

-We propose to add a sentence ”Furthermore, since the DEM used here is based on the first seven ICESat campaigns, the reference epoch will not be the same in each segment”.

- We have not considered the uncertainty of the DEM, but assumes that it is constant in each 500 m segment.
But it is clear that the variance still reflects the error of the DEM and the data, since the fit will be poorer if the errors are large.

- We agree that a weighted average should be computed to account for different number of data per repetition in one campaign. This can easily be implemented, and will be included in a revised manuscript.

3.2 Method 2

- The sentences p. 2110, l. 14-17 will be rephrased to:
”In the second method (M2), data from two ICESat campaigns are used to create a reference surface, which is used to account for topography in each segment. The reference surface is represented by a centroid point (x0,y0,H0) and slopes (dH/dx, dH/dy) and it is found by a least squares fit of these surface parameters to the measurements from two campaigns.”
The reference to Pritchard et al (2009) is removed since there are indeed significant differences between their approach and M2.

- The tracks used to generate the reference surface can be different from one segment to another.

- p. 2111, l. 8-9. By 'actual elevation change' we mean an elevation change that contributes to the mass balance estimate.

3.3 Method 3

- This method is sensitive to the track constellation in a segment. If the change in time is strongly correlated with the change in position (e.g. dH/dt and dH/dx), this method will not be able to separate the two components.

- The bar should be removed from the t in Eq. 7

3.4 Method 4.

- As mentioned earlier, we propose to skip this method in a revised manuscript, as suggested by the referee.

3.5. Elevation change results

- We agree with the referee that the variances should not be used for data editing, and we should instead derive the dV/dt from all dH/dt values and their associated variances.

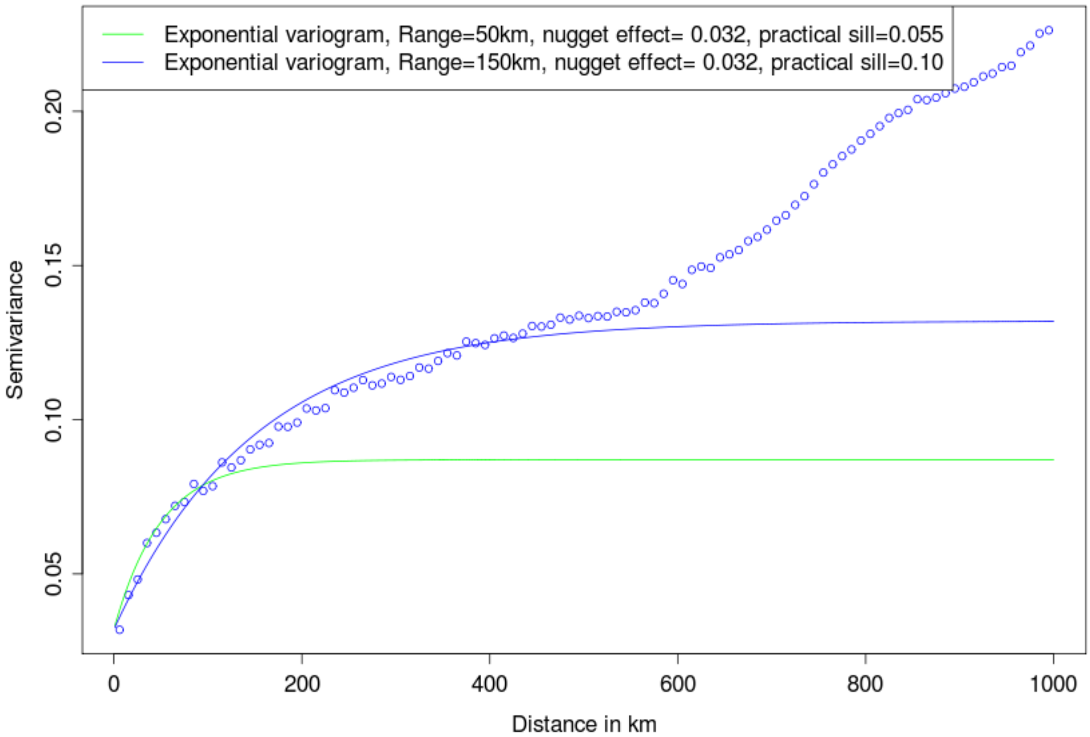
4. Deriving volume changes/4.1 Interpolation of volume changes

- Once a model of the underlying surface (the dH/dt values) is chosen, we can predict a dH/dt value at any given location and resolution. The important question is whether we believe in our model. We have done several cross validation tests to help us choose the best model (e.g. different correlations length, interpolation methods). In order to cover the entire ice sheet extrapolation will be unavoidable in some areas.
We chose, for consistency, to predict the dH/dt values at the same 5x5 km grid as the climate model and ice-mask.

-We use ordinary kriging for the interpolation. The phrase 'local neighbourhood kriging' refers to the fact that only a subset of (the closest) points are used to estimate the interpolated value.
We propose to change the sentence starting on p.2114, l. 1 ”Due to the large... .” to ”Due to the large number of dH/dt estimates, only a local subset of points is used in the kriging procedure”

-The variogram used, is based on all dH/dt data.

- We acknowledge the fact that the signal is more irregular in coastal areas, and hence the dH/dt values will vary more in these areas. But since we have all of Greenland as our study area, we have chosen to use the same variogram model everywhere. This is of cause a simplification.
Below is a plot of the experimental variogram and two variogram models. We have used the exponential variogram model with a range of 50 km (practical range of 150 km) shown with a green line on the plot.
We propose to show the variogram plot in the supplementary material of a revised manuscript.



-We have assumed isotropy, but do acknowledge that this is not the case. We have chosen to make this assumption for simplicity. We propose to add a sentence on p. 2112, l. 24:
"For simplicity it is assumed, that this variogram model is isotropic."

-It would be interesting to make a more thorough study of interpolation of satellite data, where high resolution information along-track, sparse information across-track, and area dependent variograms where taken into account. But we find that such an analysis would be outside the scope of this paper.

4.2 Bootstrapping

- Sampling with replacement does not refer to changing geographic position of the points. Sampling with replacement here refers to randomly drawing multiple times from the same pool of observed tracks.

- A resample is created in the following way: We refer to ICESat tracks of dH/dt estimates as our original data set. By use of kriging of the original data set, we get one estimate of the annual volume change (our point estimate). In order to get an uncertainty on the point estimate, N (in our case 1000) new data sets are created.. Each new bootstrapped data set is created by randomly drawing tracks from the tracks in the original data set. This sampling is done with replacement, which means that the bootstrapped data set will likely contain multiple copies of some of the original tracks. Each bootstrapped data set contains the same number of tracks as the original data set. From each of these bootstrapped data sets a new point estimate of the volume change is calculated in the same way as for the original data set. Finally the standard deviation of these N bootstrapped volume estimates is used as an estimate of the standard deviation of the original volume estimate (in a frequentist sense).

The assumption behind the bootstrap is that the observations we are sampling from are independent. Individual dH/dt points within a track can be correlated, so we consider the basic observation to be a track of dH/dt values, instead of using the individual dH/dt points. The rational is that if we want to mimic the creation of new data set, it should consist of tracks, and not isolated points.

-We will improve the explanation of the bootstrap method in a revised manuscript, by changing the following:
Paragraph (1) on p. 2114, l. 15-17 to, "Create a resample by drawing random samples with replacement. Sampling with replacement refers to, that the randomly selected observation is returned to the data pool, before a new observation is selected at random. Furthermore, it is assumed that the observations are independent. In this way a new data set is obtained with the same size as the original data set.". We propose to change the lines (23-25 on p. 2114 and 1-3 on p. 2115) to "Here, the original data set consists of the tracks of dH/dt estimates. For each method, 1000 new bootstrapped data sets are created, by randomly drawing tracks with replacement from the tracks in the original data set. Hence the bootstrapped data set will most likely contain multiple copies of some of the original tracks. Each bootstrapped data set contains the same number of tracks as the original data set. From each of these bootstrapped data sets a new estimate of the volume change is calculated in the same way as for the original data set. Finally, the standard deviation of these 1000 bootstrapped volume estimates is used as an estimate of the standard deviation of the original volume estimate (in a frequentist sense)."

4.3 Volume change results

- We agree that the method performance must be re-evaluated in a revised manuscript, after the suggested changes have been implemented.

5 . Modelling firn compaction and surface densities.

- Here, the text is misleading and the terms are not neglected. See reply to referee #2

5.3 .Interpolated metric grid

- The center figure of figure 3 is important, since it shows the original HIRHAM5 grid before interpolation.
We agree that the addition of coast lines will improve the figures, and they will be included together with the location of Station Nord.

- As the distance to the center of the HIRHAM5 grid increases, so does the error of the interpolation area. This is seen most clearly at station Nord/flade isblink in Northeast Greenland, when comparing the center and right panel.

5.5 Results of firn compaction and density modelling

- The density above the ELA is modelled as a function of surface temperature, and below the ELA ice density is assumed. Therefore, the density above the ELA is between 332 and 917 kg/m^2 depending on season and site. In Thomas et al 2006 the “firn” density is assumed to be 600 kg/m^2. To strengthen the argument of applying different densities the reference to Thomas et al. will be added to the section.

- Above the ELA the assumption is that elevation increase is due to accumulated snow/firn, whereas below the elevation increase is ice, this is to account for ice sheet dynamics and refrozen water.

- The difference between Fig. 4d and Fig. 4e, is due to the variability of both temperature and accumulation in d and only the variability of accumulation in e.

6.2 ICESat intercampaign bias correction

- We chose to follow the approach described in Gunter et al (2009) for determining the intercampaign biases. It is true that variations in MSS (in space and time) are provided along with the DNSC08 model, but these do not cover the full ICESat data time span

7. Mass balance of the GrIS

- Instead of using only the bootstrap derived error for the mass change, the different components should be added together. This will be done in at revised manuscript.
- The maximum error for the assumed densities is 37 Gt/yr, which have to be added to the final error on the measurement (see also the reply to referee #2).

8. Discussion and conclusions

- The errors are not independent of each other, and are different for each of the methods applied, this sums to the result in the discussion. However, some of the errors pointed out by the two referees and J. Bamber have to be included. Therefore, we propose to add a small section just before the discussion:
“Sum of errors.

As seen in the previous sections the error are widely different depending on the method and number of correction applied. The accumulated errors are not independent and therefore cannot be summed by addition. An estimation of the error may me done by estimating the variance of the different mass loss estimates in Table 2, which results in a 2σ error of the different estimates of 44.9. “

- Such a figure/table is already presented in the SWIPA report [AMAP 2009] and a reference will be added.

- It is pointed out by the referee that it does not make sense to compare our mass change result with the one from Slobbe et al 2009, due to different approaches and time span. We find this a bit confusing, because following this argument it would not make sense to compare to any studies which have used different data, time span and approaches?.
We propose to rephrase the sentences p. 2128 l. 1-5 to:
”Finally, our total mass balance result is large compared to the ICESat derived mass

loss of 139±68 Gt yr−1 found by Slobbe et al. (2009), based on data from 2003 to 2007, and which does not include firn compaction, elastic uplift and intercampaign bias corrections.
We believe that we have presented a novel approach to the application of ICESat data in estimating the total mass balance of the GrIS, by including firn compaction and density modelling."

3. Technical corrections

-p. 2105, l. 10-11: Will be rephrased to:
"Firn dynamics and surface densities are important factors in deriving the mass change from remote sensing altimetry."

- p.2105, l. 19-21 Will be rephrased to:
"This result is in good agreement with several other studies of the Greenland ice sheet mass balance, based on different remote sensing techniques"

-p. 2107 this sentence will be added:
"In Sect. 6, elevation changes that are not contributing to the mass balance of the ice sheet are described and quantified."

-p. 2107, l. 8-9 the sentence will be rephrased to:
"As a consequence of this and due to the inclination of the satellite, successive tracks are separated by approximately 30 km in the southern part of Greenland."

-p. 2108, l.15-15 The sentence will be rephrased to:
" Thus besides seasonal variations and secular trends, the observed elevation difference between tracks contains a contribution from the terrain."

- p. 2117, l. 14. Revision of equation 10, as described in the answer to referee #2, clarifies the symbol used fore mass balance.

- Caption Fig. 4. "the" will be removed.

References:

Bea Csatho, Toni. Schenk, C.J. van der Veen, William B Krabill, http://websrv.cs.umt.edu/isis/index.php/Present_Day_Greenland, Presented at the AGU 2009 Fall Meeting

Smith, B. E., Fricker, H. A., Joughin, I. R. and Tulaczyk, S. : An inventory of active subglacial lakes in Antarctica detected by ICESat (2003-2008), Journal of Glaciology, vol 55, 2009.

Gunter, B., Urban, T., Riva, R., Helsen, M., Harpold, R., Poole, S., Nagel, P., Schutz, B., and Tapley, B.: A comparison of coincident GRACE and ICESat data over Antarctica, J. Geodesy, 83, 1051–1060,

Davison, A. C. and Hinkley, D.: Bootstrap Methods and their Application, 8th edn., Cambridge

Series in Statistical and Probabilistic Mathematics, Meteorological Institute, University of Bonn, Bonn, 2006. 2114

AMAP 2009 Dahl-Jensen, D.; Bamber, J.; Bøggild, C.; Buch, E.; Christensen, J.; Dethloff, K.; M. Fahnestock; Marshall, S.; Rosing, M.; Steffen, K.; Thomas, R.; Truffer, M.; van den Broeke, M. & C.J. van der Veen. (AMAP, 2009).The Greenland Ice Sheet in a Changing Climate: Snow, Water, Ice and Permafrost in the Arctic (SWIPA) 2009, *Arctic Monitoring and Assessment Programme (AMAP), Oslo.*

Thomas, R., Frederick, E., Krabill, W., Manizade, S. Martin, C., 2006. Progressive increase in ice loss from Greenland, Geophys. Res. Lett., 33, L10503.

Slobbe, D., Ditmar, P., and Lindenbergh, R.: Estimating the rates of mass change, ice volume change and snow volume change in Greenland from ICESat and GRACE data, Geophys. J. Int., 176, 95–106, doi:10.1111/j.1365-246X.2008.03978.x, 2009. 2128