Authors response to the comments of referee #1

We thank Johan Nilsson for the thorough, fast and helpful evaluation of the manuscript. Especially the comments concerning the additional IceBridge dataset helped to get further confidence in the absolute accuracy of the results. Also the smaller edit recommendations in the supplement helped us a lot to further clarify some statements and improve the language. We would like to mention that the correction of a small inconsistency in our variance propagation of the GNSS profiles and the variable ICESat campaign precisions (recommended by reviewer #2) slightly changed the results for the campaign biases. However, this does not change anything in the general messages of the manuscript. In the following we will respond the comments one by one.

(1) One can clearly see that the inter-campaign biases differ from each other depending on the surface type or region used to derive them. What is the sensitivity of your or other solutions to surface type or possible location bias? I would like, if possible, some more discussion about this. This as the estimation of the inter-campaign bias I find to be the main outcome of this study.

We agree that inter-comparisons of ICESat biases obtained with different methods and in different regions are needed, although this is beyond the scope of the present paper. Towards the application of this set of biases in other glaciated regions, especially in Greenland, further investigations should be carried out. Assuming that these biases are laser-energy-related (Schutz et al., 2011) those corrections would be applicable for surfaces with similar albedos (Hofton et al., 2013) as Lake Vostok.

(2) The choice of the crossover methods for the validation procedure, though very accurate and mature, has the limitations of limited spatial coverage and data density. It would be useful to include, or at least compare, the use of another method to judge the stability of the distribution used to derive the statistics. A small-scale study using the "average footprint method", which was initially discussed in the paper and applied in Wouters et al. (2015), might be of interest?

Wouters et al. (2015) validate the CryoSat-2 elevation change grid using Ice-Brigde ATM derived dh/dt, gridded to a similar resolution. However, ATM measures a swath of elevations and dh/dt can be considered uniform over longer distances (e.g. a beam limited radar altimeter footprint), which is not the case for the elevation itself. Thus, for computing elevation grids, we would need not just linear profiles but a good 2-D coverage, preferably an array of profiles (Phillips et al., 1998). Our dataset does not include such arrays and thus does not provide for such an analysis, although it would indeed be very interesting.

(3) The application of the 67 cm elevation bias in the CryoSat-2 Baseline C products I think needs to be discussed in a bit more detail, as this is one of the main differences between the two ESA baselines. One can clearly see that the application of the bias pushes the radar horizon closer to the reference surface. For full-waveform retrackers, like ESA's, this has no major effect as they usually show high surface penetration. However, leading edge retrackers, which track closer to the surface (not using the entire waveform), will in many cases produce positive biases. This could, for example, be seen in your presentation at ESA living planet. As it forces the radar-laser surface bias to be positive, which is unphysical for radars, it begs the question: should the bias be applied?

On Lake Vostok, there are no positive biases for any dataset. The cause for the shift of the biases towards the positive direction has been explained for the Envisat data and is exactly the same for CryoSat-2 LRM. It is the sampling of the local topography maxima by the POCAs (point of closest approach) of the radar altimeter, compared to the full topography sampling for the kinematic data. Thus, with increasing slope and increasing roughness too, the biases grow towards positive. This is discussed in section 3.3.1.

(4) Stated in the manuscript is the preference for threshold retrackers over model based ones, as they are less affected by volume scattering, which has also been proven in other studies. However, I can't find any information about which threshold is used for the OCOG retracker for Envisat (25% for CryoSat-2)? It would be good to state them clearly in the manuscript.

Done. This information was very difficult to find. It is defined in the auxiliary processing file 'RA2_ICT_AX: Ice1/Sea Ice Configuration'. Finally, with the help of the ESA support, we found out that the threshold is set to 30%.

(5) Figure 6 shows the relation between the relative change in elevation and backscatter at crossover locations over Lake Vostok. It would be interesting to see the effect of other waveform parameters, like LeW and TeS, in the same type of plot (preferably for both Envisat and CryoSat-2). Remy et al. (2012) suggested that the addition of these parameters would be more valuable understanding snow characteristic fluctuations using Envisat and CryoSat-2 derived time series. These parameters are already available or can be easily computed for the two products. I think the inclusion of these parameters in your analysis would provide heavier weight to your argument.

This would indeed be interesting, but it goes beyond the scope of this validation. What we want to show is, that the seasonal variation is an order of magnitude larger for functional fits and that this signal is highly correlated to backscatter. As our final recommendation is to use low-threshold retrackers (where no LeW and TeS exist) we left those parameters out of our analysis.

(6) I might be misunderstanding you, but what is the crossover time-span used to derive your validation statistics? I'm guessing you use all crossovers independent of time-span, as your elevation change signal is very small? If so, it would be good to state this in more detail in the manuscript, as crossover difference of less than 30 days has been the norm for previous validation studies, see (Khvorostovsky 2012).

You are guessing right. Except for the validation of ICESat, we do assume that the surface is stable and thus, that the elevation change is negligible. This is described more in detail now in Sect. 2.5.

(7) The Bernese software has been used for the processing of the GNSS data. How does this software compare to the other available packages, like GIPSY? I think it would be good to discuss the impact (or differences) of different software packages on the results, maybe in sentence or two? Further, maybe direct the reader to a set of references for the curious.

Such a survey has been performed by Dietrich et al. (2001) for static observations. In a comparison of the results from six groups using four different software packages (including Bernese and GIPSY) the results showed a very good agreement. However, we would like to emphasize, that even within the same software package the results strongly depend on the strategy, the algorithms and the models and products used. Kohler et al. (2013) use Precise Point Positioning (PPP) which does not require reference stations but instead needs precise satellite clock information for every epoch. Usually these clock offset data have a rate of 30s. Prior to 2004, the rates were even more sparse (5min). Only after 2008, the Center for Orbit Determination in Europe (CODE) publishes a 5s high rate product but this would not be applicable for the early profiles. Here, the significantly decreasing number of observations would lead to much larger errors in a PPP-solution.

However, our results are very similar to the results of Kohler et al. (2013). They process two seasons with two profiles each. From crossovers between these profiles they calculate a mean offset of -4.1 cm and a standard deviation of 8.6 cm. We divide our average crossover difference by $\sqrt{(2)}$ to obtain our accuracy measure $\overline{RMS_X}$, which is in the range from 4–9 cm. Thus we conclude that the accuracy and precision of both types of solutions is very similar.

(8) Has there been any attempt to validate or compare elevations over Lake Vostok using NASA's Operation IceBridge, to acquire another reference surface? If so how do they compare?

Very good point! Has been done in section 2.4 now and included in Tab. 1.

(9) The difference between the Bamber et al. (2009) DEM and Bedmap-2 is very interesting! Clearly the rounding will have an effect on the precision, but the -1 m bias is surprising? Is this bias spatially dependent, as different datasets have been merged together?

This bias exists all over Antarctica where the Bamber-DEM is the data source for Bedmap-2. It varies mainly between -0.5 to -1.5 m along bands which look like waves. Hopefully this bug will be found and fixes in case of a new version of Bedmap.

References

Dietrich, R., Dach, R., Engelhardt, G., Ihde, J., Korth, W., Kutterer, H.-J., Lindner, K., Mayer, M., Menge, F., Miller, H., Müller, C., Niemeier, W., Perlt, J., Pohl, M., Salbach, H., Schenke, H.-W., Schöne, T., Seeber, G., Veit, A., and Völksen, C.: ITRF coordinates and plate velocities from repeated GPS campaigns in Antarctica – an analysis based on different individual solutions, J. Geod., 74, 756–766, 2001.

Hofton, M., Luthcke, S., and Blair, J.: Estimation of ICESat intercampaign elevation biases from comparison of lidar data in East Antarctica: Estimating ICESat elevation biases, Geophys. Res. Lett., 40, 5698–5703, doi: 10.1002/2013GL057652, 2013.

Kohler, J., Neumann, T., Robbins, J., Tronstad, S., and Melland, G.: ICESat Elevations in Antarctica Along the 2007–09 Norway-USA Traverse: Validation With Ground-Based GPS, IEEE Trans. Geosci. Remote Sens., 51, 1578–1587, doi:10.1109/TGRS.2012.2207963, 2013.

Phillips, H., I., A., Coleman, R., Hyland, G., Morgan, P., and Young, N.: Comparison of ERS satellite radar altimeter heights with GPS-derived heights on the Amery Ice Shelf, East Antarctica, Ann. Glac., 27, 19–24, 1998.

Schutz, B., DiMarzio, J., Luthcke, S., Hancock, D., and Urban, T.: Notice Concerning Detection of ICESat/GLAS Inter-Campaign Elevation Biases, Tech. rep., 2011.

Wouters, B., Martin-Español, A., Helm, V., Flament, T., van Wessem, J., Ligtenberg, S., van den Broeke, M., and Bamber, J.: Dynamic thinning of glaciers on the Southern Antarctic Peninsula, Science, 348, 899–903, doi: 10.1126/science.aaa5727, 2015.