Reviewer#2

This paper presents an impressive and consistent new set of elevation change data for a significant amount of glaciers in the Himalaya-Karakoram-Pamir region based on a comparison of SRTM and SPOT digital elevation models. Thus, the authors provide a sound estimate of glacier mass balances (including their spatial variability) for a highly debated region. I fully agree with the first reviewer that this study is well performed, the methods and results are clearly described and that the article deserves to be rapidly published. Nevertheless, I have some more substantive comments that might require some additional discussion in a second version of the paper. These are not meant to criticize the presented results but might help the authors to refine some of their conclusions.

Substantive comments

- **Radar penetration correction**

The estimation of radar penetration depth by comparison of the different frequency bands is reasonable. However, it would be helpful to already provide the order of magnitude of this correction on page 985 (i.e. in the method description) to allow a judgment of its importance. This would be better than just referring to Gardelle et al. (2012b). Furthermore, I am not sure if it is given than the X-band (9.7 GHz) has no penetration depth at all. Obviously it is less than for the 5.7 GHz band, but the total effect is likely to be rather under- than overestimated with the correction. Wouldn't it be possible to dig deeper into this issue using GPR theory?

> Because Table 3 will be placed somewhere close to the paragraph, we do not think it is necessary to provide the order of magnitude of the penetration. We already stated in the text that it is of "several meters", probably sufficient.
>
> We agree with the reviewer that the non-penetration of the X-band has yet to be confirmed. It was clearly stated in (Gardelle et al., 2012): "*Clearly, this hypothesis requires further validation, especially by the German Aerospace Center (DLR) TanDEM-X mission*". In the revised MS, to make sure the reader is aware of this strong hypothesis, we write "*Since the X-band penetration is expected to be low compared to C-band penetration (an hypothesis that still needs to be confirmed; Ulaby et al, 1986), …*". The relevance of the SRTM penetration measured regionally for a specific glacier is also discussed now in the revised MS, section 5.1.
>
> We have now modify the calculation of the error estimate regarding this correction (see general responses)
>
> We do not think the present paper is the place to dig deeper into GPR theory. Part of the effort to dig into the GPR literature was made by (Gardelle et al., 2012) but was not really conclusive. Constraining the radar penetration using GPR theory may actually require some data (e.g., temperature, density profile, humidity content of the snowpack) that are hardly available at the time of the SRTM mission, especially in PKH.

- **Seasonality correction**

The quantification of the winter accumulation rates used for the seasonality correction is weak. One single value for the Karakoram is available from the 1980s. For all other glaciers the mean of surveyed glaciers in the Northern Hemisphere is used, i.e. winter accumulations in the Himalaya-Pamir region are quantified by including e.g. maritime glaciers in Norway... A better estimate could probably be achieved using almost every method (analysis of precipitation data, isolated data on accumulation rates from ice cores, etc). For glaciers with a summer-accumulation type (Himalaya), the winter accumulation rates are probably overestimated. The authors cover the large uncertainties in this correction with their error bars. Nevertheless, I suggest to try and get a more reasonable estimate that takes into account local characteristics.

See General response 0.2 where we built on the reviewer's comment to apply a different seasonality correction in the East and in the West.

We disagree with the statement that "better estimate could probably be achieved using almost every method". Precipitations are notoriously difficult to measure in the mountains and measurements in the valleys may severely underestimate (by a factor of 2-3, probably more when the stations are located in the dry valleys of the Karakoram) the accumulation on glaciers if not corrected for a realistic accumulation gradient with altitude that is glacier-specific (Immerzeel et al., 2012; Vincent, 2002). Furthermore, a single accumulation measurement (for example at an ice core drill site on a single glacier) even if available may not reflect the glacier-wide accumulation because of its high spatial variability (Azam et al., submitted; Machguth et al., 2006).

- **Off-glacier elevation changes**

Obviously, the off-glacier elevation changes between the SRTM and SPOT DEMs are an excellent mean to quantify the uncertainties. Numbers are provided by the authors (page 988, line 16) but are not further discussed. It would be highly beneficial to go into some more details here: Do the off-glacier elevation changes show an elevation dependence? Are there some significant differences between the study regions that might indicate regional biases? Are the off-glacier elevation changes equally distributed within one scene, i.e. are they consistent between the center and the edges of the scene? This discussion might be valuable to judge the spatial representativeness of the error bars.

In the revised MS, we have now:

(i) added a histogram showing the distribution of the elevation change off glaciers on each map of elevation change. By definition the mean difference is 0. The median and the standard deviation of the elevation difference are also given for each site.

(ii) shown the map of the elevation changes off glaciers in a supplement so that the readers can verify that there are very small spatial variations in the bias off glaciers (typically less then 1 m at length scale of a few kilometers). This is now briefly described at the start of sub-section 4.1.

Regarding a possible elevation dependence of the bias, this issue has been examined in detail in a previous study (Gardelle et al., 2012) where we proposed a specific correction which is used here and whose aim is to remove this possible bias.

- **Comparison of discharge to Kaser et al. (2010)**

The authors calculate runoff contributions due to glacier imbalance and compare these numbers to observed runoff in the main streams draining the study region. They acknowledge that only annual contributions can be quantified, and refer to Kaser et al. (2010) for seasonal contributions. I am troubled by this comparison and do not think that it is possible: Kaser et al. (2010) have based their analysis on global climate data sets but do neither include direct data on glacier mass balance nor runoff. The approach – and also the results – are thus inconsistent with the percentage contributions presented here. This probably explains the somewhat strange numbers given in Table 6: How can the seasonal glacier contribution (most probably the authors refer to the melt season here, although it is not stated) be smaller (!) than the annual mean contribution? This would require a better discussion of the results by Kaser et al. and more details on their approach, but I would just suggest to remove the comparison here as the methodologies are different and the numbers are rather worrying than helpful.

> This section of our paper was not clear enough and, thus, perhaps misunderstood. Importantly, we failed to mention that the Kaser et al. (2010) analysis was based on the assumption that glacier mass budgets are in equilibrium. This is important to mention because it allows comparison of their values with ours. This assumption also explains why the seasonally-delayed contribution can be equal to 0 (end-member case of a summer-type accumulation glacier where, every month, the accumulated snow melts away) and still, the imbalance contribution be different from 0. Thus, Kaser et al. (2010) and our study are measuring two different, non-overlapping, components of the glacier contribution to river discharge. The methods to measure these two components are fundamentally different (by necessity) but it does not mean that the final numbers cannot be compared. We hope that, with our improved description of Kaser et al. (2010), this section will be better understood. We stress however that our paper is not the place to describe in full detail the methods of Kaser et al. (2010). Overall, we believe the comparison of seasonal contributions and imbalance contributions can add important insight to the current discussion of glacier contribution to river runoff by separating the seasonal contribution (not related to climate change) from the (climate-change related) imbalance contribution. Due to the difficulty of summarizing these details in just 1-2 sentences, the decadal glacier mass loss contribution to river discharge is not mention anymore in the abstract.

- **Thinning over debris-covered ice**

The authors convincingly show that the thinning of debris-covered ice is not smaller compared to clean ice. This would be expected from the well-documented melt reduction below supraglacial debris. The authors interpret this observation with differences in ice dynamics. I have the impression that the comparison of elevation change rates over debris-covered and clean ice surfaces might be biased (explanations see below). Based on my comments the authors might consider adding some more discussion on this important issue. Surface elevation change rates at given altitudes within individual regions are performed. The approach of comparing identical altitudes only is sound and removes a possible elevation bias. However, do the authors also consider *glacier size / elevation range* in their evaluation? In my opinion, a direct comparison of elevation changes over debris-covered and debris-free surfaces is only feasible for glaciers that exhibit the same elevation range, and thus comparable ice flow dynamics. I would speculate that clean ice surfaces at low elevations (e.g. below 4000 m a.s.l.) are just found on smaller glaciers, i.e. glaciers with relatively high accumulation rates and low ELAs, and that almost all glaciers with a large elevation range have debris-covered tongues. This might lead to completely different dynamic responses of the two glacier types to climatic changes that make an immediate comparison of the dH/dt impossible. Furthermore, the statistical representativeness would also need to be discussed: How many data points for clean ice are available at low elevation in comparison to debris-covered pixels?

First, in the submitted MS, we did not interpret the differential thinning rates only in term of ice dynamics (if it appeared so then our writing did not reflect our understanding). We did not discard the fact that surface ablation could actually be similar or higher on debris covered surface due to some features, such as ice cliffs or lakes, enhancing ablation even when the debris cover is thick. We tried to make this clearer in the revised MS and, building upon G. Cogley's comments and a recent paper (Zhang et al., 2013b), we now discuss that those thinning rates can be explained by a debris cover which is, on average, thinner than the thickness threshold between enhanced/reduced ablation.

Furthermore, in the ICESat study of Kaab et al. (2012) similar thinning rates over clean and debris-covered ice are found by comparing neighboring pixels (average distance between them is 1 km), most of them likely located on the same glacier and thus with probably similar dynamic, on average.

The Everest study site, where thinning is higher under debris (a confirmation of a previous study by (Nuimura et al., 2012) who had not performed a histogram adjustment though) would be a very interesting place to examine in more detail the response of individual glaciers and the relationship between thinning rate and the % of debris coverage, the altitude range, the size of the glaciers and the role of differential ice dynamics. But we believe that this glacier-by-glacier analysis is beyond the scope of our analysis that has a regional focus.

Regarding statistical representativeness, we show below the number of pixels in the lowest elevation bin. Note also that the number of pixels is rapidly increasing when elevation is increasing. The total differs but the fact that, generally, >100 pixels are present in this lowest elevation band give us some confidences in the differences observed.

	Pamir	Karak. west	Karak. east	Spiti	West Nepal	Everest	Bhutan	Hengduan Shan
Lowest elevation	2900-3000	3000-3100	3300-3400	4200-4300	4500-4600	4400-4500	4300-4400	3300-3400
Nb. debris	1361	1190	196	279	818	1975	2183	716
Nb clean	81	192	123	132	111	553	142	125

- **Consideration of year-to-year mass balance variability:**

The study provides an extensive validation of calculated mass changes against previous studies. However, I miss a comparison to direct glaciological time series. I am aware that very little is available for the region and that the uncertainties are high. Nevertheless, annual mass balance time series (such as from Chhota Shigri Glacier, Azam et al., 2012) might provide some valuable information about year-to-year variability. Strictly speaking the validation of the period mean mass balances with other studies (covering slightly different periods) is only possible after removing artefacts coming from year-to-year variability. It is impossible to provide a sound correction based on the available in-situ mass balance data sets but it would be interesting to see a short discussion about the mass balance variability within the considered 11-year period and whether this variability might explain some of the disagreement with other studies. The present results mostly give smaller mass losses (Fig. 5, Table A1). Could this observation simply be explained by above average mass balance in the last years (i.e. after about 2008) that are covered by this, but not by the other studies (Bolch et al., 2011; Nuimura et al., 2012; Kaaeb et al., 2012; Berthier et al., 2007)?

By definition, the geodetic method does not provide the mass balance variability during the study period, but only the cumulative mass balance. As stated by the reviewer, it would be illusive to attempt a region-wide correction of the mass balances measured over different periods to make them exactly comparable. Following the reviewer's advice we have now included a full paragraph to make clear that mass balances should be compared with care when they do not span the same time period. To our knowledge, Chhota Shigri Glacier is the only peer-reviewed (Azam et al., 2012; Vincent et al., 2013) annual mass balance record that can be used to estimate this year-to-year variability for most of the first decade of the 21st century. The inter-annual variability from this record is now quoted in the text, together with the one from Abramov Glacier (WGMS, 2012). We also justify in the revised MS why we do not compare with more field mass balance records.

Specific comments

• page 976, line 15-17: This sentence is difficult to understand in the abstract. Following my substantive comment above, I recommend omitting it or replacing it with another important conclusion.

We omitted the sentence and now concentrated the last two sentence of the abstract on a comparison to global glacier mass balance.

• page 977, line 15: A short definition of the glacier imbalance in the present context would be helpful.

Reworded to avoid "imbalance", "decacal mass loss" is used instead.

• page 984, line 3: The ELA digitized from Landsat images corresponds to this one given year and might show a considerable variability. This might need to be acknowledged in a sentence and/or some references could be provided to back up the assumption of a constant ELA.

We agree that this is a strong assumption and that ideally regional ELA should be measured throughout the study period. This is now better reflected in the text (see revised section 3.2). As suggested by T. Nuimura, we added the standard deviation of those ELAs.

• page 986, line 13-15: The seasonality correction would only be lower by 1-2 orders of magnitude than the cumulative signal if the mass balances are significantly different from zero. With the balanced conditions in the Karakoram and the Pamir the uncertainty in this correction might well make the difference between a positive and a negative mass budget.

True, our statement was only valid for the study sites where the mass balance is clearly negative. The statement has been removed.

• page 987, line 16: The density assumption might require some more discussion as it linearly influences the final results. Will the density of volume change be the same for all regions although they exhibit strongly different mass balances?

> We now describe two alternative density scenarios (i) 900 kg/m^3 everywhere (Sorge's law) and (ii) 600 kg/m^3 in the accumulation area and 900 kg/m^3 in the ablation area. The same two density scenarios where used in Kaab et al. (2012) and in earlier papers. We have estimated the maximum difference between our preferred scenario (850±60 kg/m3 everywhere) and those two others scenarios. For the eastern sites with negative mass balances, the maximum difference is small, at 0.03 m w.e./yr. For the western sites, the maximum difference is higher, at 0.06 m w.e./yr. Those uncertainties, due to the choice of a given density scenario, remain low compared to other sources of errors. Those alternative density scenarios are now included in the revised text (section 3.5)

• page 990, line 6: Elevation changes averaged over the ablation area might be mistaken as mass balances / melt rates. I see the benefit of discussing these data here but I would suggest to clearly state the meaning of ablation area elevation changes and their limitations.

> This section is now shorter and discusses rates of elevation changes in the two zones to avoid that the reader interprets them as melt rates. We have also tried to better highlight some of the important pattern revealed by these maps (e.g., no elevation changes in the accumulation zone of all the eastern study sites).

• page 991, line 25: Are there any explanations for these strong differences in the mass balance of neighbouring glaciers? Whereas the authors discuss mass balance differences between the regions in detail (in connection with climatic patterns) the glacier-to-glacier variation in mass balance (which can obviously be significant) is not addressed.

> First, we stress that we did not compute the mass balance for each individual glaciers on each study site but simply singularized out some emblematic glaciers (Siachen, Fedtchenko, Baltoro, Rongbuk), some glaciers that are followed in the field or some glaciers previously observed using the geodetic method (e.g., in the Everest area). Examining systematically glacier-to-glacier mass balance variability would require an additional, non trivial, step which consist in splitting the inventory in individual glaciers. A multiple regression between glacier mass balance and topographic parameters/debris coverage such as perform by Huss (2012) would certainly be an interesting next study. But we think that this is beyond the scope of the present paper that is dedicated to regional assessment.

• page 993, line 22: Interesting. Can these numbers be put into physical context? i.e. do they correspond to a penetration into the winter snow coverage only, or also into the uppermost firn layers?

> Penetration can reach up to 7-8 m in the upper reaches of Karakoram for example, so the radar signal probably go deeper that the annual snow layer. However, given that there is a complete lack of knowledge of the winter accumulation for most (if not all) these study sites we do not think it is possible to really put these numbers into a physical context. To do so, it would actually require some measurements (snow pits and if possible more than one given the 90 m SRTM pixel size) in mid-February 2000 at the time of the SRTM mission.

• page 997, line 8: Here and elsewhere: unit: m yr−1 w.e. Wouldn't it be more logical to write it as m w.e. yr−1 ?

Changed everywhere. Indeed more logical.

• page 999, line 8-22: Although interesting I was not quite sure if this paragraph is actually necessary for the results / conclusions of this paper. The topic is only loosely related and it could be removed.

Agreed and deleted (the same suggestion was made by T. Nuimura).

• page 1001, line 4: I think, most importantly high-elevation precipitation measurements would be needed. And weather stations in these environments probably have troubles in accurately determining precipitation. So, a sentence might be added that direct measurements of accumulation on High Mountain Asia glaciers would (also) be required to understand to ongoing processes.

We fully agree with the referee. The statement was changed and a reference to (Azam et al., submitted) is added as an example of the sort of studies that are needed to better estimate annual accumulation and understand interannual/decadal trend in both summer and winter mass balance.