*Thanks to both reviewers and the editor for such careful and constructive comments. They are much appreciated! We respond to comments point-by-point in blue italics, and have used red text in the manuscript to indicate revisions.*

18-Dec-2017   
  
Dear Ms. Pulwicki,   
  
JOG-17-0123 entitled "Uncertainties in estimating winter balance from direct measurements of snow depth and density on alpine glaciers" which you submitted to the Journal of Glaciology, has been reviewed.  The comments of the reviewer(s) and editor are included at the bottom of this letter.   
  
While the reviewer(s) and editor recognise the potential of your manuscript as a valuable contribution to the journal, they also suggest some major revisions to your manuscript. Therefore, I invite you to respond to their comments and revise your manuscript.   
  
To start your revision now, click the link below:   
  
\*\*\* PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. \*\*\*   
  
https://mc.manuscriptcentral.com/jog?URL\_MASK=c4f6bc6939974504b4e71a0271d844e7   
  
Alternatively, you may log into your Author Centre at https://mc.manuscriptcentral.com/jog, where you will find your manuscript under "Manuscripts Awaiting Revision". Upon submission of the revised version, your manuscript number will be appended to denote a revision.   
  
When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) in the space provided.  Please use this space to document any changes you make to the original manuscript.  In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s).   
  
Because we are trying to facilitate timely publication of manuscripts submitted to the Journal of Glaciology, your revised manuscript should be uploaded as soon as possible. We expect to receive your revision by 17-Jan-2018. If it is not possible for you to submit your revision by this date, please contact the Editorial Office to rearrange the due date. Otherwise we may have to consider your paper as a new submission.   
  
Once again, thank you for submitting your manuscript to the Journal of Glaciology and I look forward to receiving your revision.   
  
Best regards,   
Dr. Nicolas Eckert   
Scientific Editor, Journal of Glaciology   
nicolas.eckert@irstea.fr   
  
Scientific Editor: Eckert, Nicolas   
Comments to the Author:   
(There are no comments.) 

Reviewer: 1   
  
Comments to the Author   
General comments to the authors   
  
This paper presents a large number of winter balance measurements over three glaciers. Through the statistical analysis of these measurements, the study estimates glacier-wide winter mass balance and the associated uncertainties. This study is a valuable contribution since understanding the uncertainties and modeling winter mass balance variability remains challenge within the scientific community. The large number of measurements is impressive and an important contribution in itself. With some minor revisions, I recommend that this study for publication.   
  
This study is very thorough and contains a lot of information. While this information is useful, it the main objective is often lost in the details. I would suggest reorganizing the structure and splitting the “Result and Discussion” section in two distinct sections. For instance, the result part could include only the glacier-wide mass balance computed with LR and KG methods based on data measurements with the associated uncertainties and the comparison of these two approaches. This seems to be the main objective of the study. The second part of the discussion would include the spatial variably of winter balance at glacier scale (with the influence of topographic parameters), as well as the difference between glaciers and a discussion on the complexity of obtaining a transferable relationship, as mentioned in this paper. The regional winter balance gradient could also be mentioned here. The section “Limitations and future work” could also be included in the second part of the discussion. Please, refer to my comments below for more details.   
  
My second general comment is that the method used to calculate the LR is interesting but complex. Thus this approach could be explained in more detail (by providing a scheme and/or some equations).   
  
Regarding the formatting, each section should be assigned a section number.

*We used the LaTex formatting provided by IGS and it did not include section titles. Not sure how to add them…*  
  
Detailed comments   
  
Introduction   
- Line 33: Please add a reference

“Winter balance (WB) is notoriously difficult to estimate *(e.g. Dadi ́c and others, 2010; Cogley and others, 2011)* “

- Lines 34 to 39: Are there the only reason? Isn’t it also because the access on these glaciers during the winter period is more difficult?

“Simultaneously extensive, high resolution and accurate snow distribution measurements on glaciers are therefore difficult to obtain (e.g. Cogley and others, 2011; McGrath and others, 2015) *and is further complicated by the inaccessibility of many glacierized regions during the winter.*”

- Lines 62-63: A reference to the work made by Cullen et al., (2017) based on a co-kriging method could be mentioned here.

“…as well as geospatial interpolation (e.g. Erxleben and others, 2002; *Cullen and others, 2017)* including various forms of kriging.”

- Line 64: The SnowTran-3D model (Liston and Sturm) could be mentioned here.

“Physical snow models such as *SnowTran-3D (Liston and Sturm, 1998),* Alpine3D (Lehning and others, 2006), and SnowDrift3D(Schneiderbauer and Prokop, 2011) are widely used,…”

Study Site   
- Fig.1: Grey squares are not visible. Please change color or/and extend size.   
- Table 1: It is surprising that glacier 2 and glacier 4 have the same ELA as there are oriented in the opposite direction!   
It will be interesting and useful for future discussions to add in this Table the elevation range of the measurements, and the % of the glacier covered by these measurements (e.g. using the % of grid cell of the DEM containing at least one measurement or something similar).   
  
Methods   
- Line 102: To facilitate the reading, add a paragraph (before “sampling design”) with general information (mentioned below) about the field campaign (e.g. Line 119, lines 122-123, see my comments below).   
- Line 114: Is there a reference for this method?   
- Line 115: Fig. 1f (not 1e)   
- Line 119: This is not only about the snow depth. You could add this sentence in the “introduction paragraph” of the field measurements.   
- Lines 122-124: It is also about the campaign in general. Add this in the “intro paragraph”.   
- Lines 135-136: There is here mixing information between snow depth and density. Please mention here only the snow depth measurement method as the density method is explained below.   
- Lines 138-139: This has already been said in the section “Sampling design”. So you can delete this sentence and add the reference “Shea and Jamieson, 2010” on line 114.   
- Lines 143-144: Is there a specific reason for this fourth zigzag measurement?   
- Line 146: This is quite confusing. I suggest to delete “as well as…Federal Sampler”; or add the number of density measurements made using this method, and start the next sentence (snow pit description) in a new paragraph to make it clearer.   
- Line 146: Please define the acronym (SP) here (i.e. the first time that you mention it) and then, only use the acronym (e.g. line 147, 158…) to be consistent.   
- Line 147: Same remark for Federal Sampler acronym (SF).   
- Lines 149-155: To be consistent, if the density measurement uncertainty is mentioned here, the snow depth measurement uncertainty should also be quantified in the previous section (i.e. uncertainty due to probe and snow pits methods).   
- Line 165: FS method uncertainty should also be evaluated to remain consistent.   
- Lines 166-172: This is general information about the field campaign, and concern also the snow depth. This paragraph could be in the “introduction paragraph” mentioned above.   
- Lines 171-172: It could be evaluated using a modeling approach but this is not the aim of the study. So you could at least provide an estimation of the uncertainty related to these events.   
- Line 180: These acronyms have already been defined before. So please use only the acronym.   
- Line 183: Information about DEM is only available in the supplementary material. Please mention here the year and the source of the DEM.   
- Line 189: The uncertainty that you mention here is not clear. The uncertainty concerns the snow depth difference between two observers distant from 10m, or is it the difference in measurement when the next observer arrives at the location of the one before him? And how the uncertainty is evaluated?   
- Line 191: Can you quantify this difference? Is it taken into account?   
- Line 197: Acronym SK already defined line 194. So please use only the acronym here and above (e.g. lines 224, 226, 232…)   
- Line 194: “Regression” or “Linear regression”, please be consistent in the article and use the acronym once defined.   
- Line 201: The wind-redistribution parameter should be defined (and not only in the supplementary material). Please add “Sx” and the reference to Winstral and Marks’study.   
- Lines 207-208: Refer to my comment on the Supplementary material.   
- Lines 199-222: This is an interesting and complete method which have been used here, but also quite complex. This paragraph mentions many references which need to be read to properly understand the method. I suggest providing more information and detail about the approach used. Some equations and/or a scheme could help the understanding. I think this is important, as it is one of the main objectives of this study.   
- Line 224: Use only the acronym.   
- Line 232: “topographic regression” has not been defined, even if we can guess what is it. Please be consistent when using “Linear regression”, “regression” and “topographic regression” after defined it to avoid confusions.   
- Lines 232-234: So why this method has been applied? Please explain and justify why you choose to use this approach.   
- Line 235: The specific section “uncertainty analysis” is confusing. In reading the article we expect to find all information about uncertainties in this section (including an evaluation of measurements uncertainties). I suggest renaming the section (e.g. “Uncertainty evaluation on glacier-wide WB” or “Uncertainty analysis using a Monte Carlo approach”, or something similar).   
  
Results and discussion   
- Fig 2.: Should be in this section, not before.   
Fig. 2b: Why the point G13\_ASP doesn’t have an x-error bar?   
In the legend, change “Labels indicate snow pit locations” by “Labels indicate SP locations”   
- Line 278: Add “a” as following: “Fig. 2a”.   
- Fig. 3: Please use more contrasted colors. For instance, it is almost impossible to see the difference between M2 and U in Fig 3c.   
- Lines 288, 291,297,302,307… Change Feredal Sampler by FS.   
- Line 289: The diameter size information should be in the method section.   
- Line 302-303: Can you be more specific and provide more information regarding the “the conditions at the time of the sampling”.   
- Line 311: Please justify why do you use the FS density measurements, while you said that there is a bias that you cannot correct.   
- Line 323: Where are these results? Are they “not shown”? As you have a supplementary material, it could be used to insert these results.  
- Lines 324-326: Same remark than above.   
- Line 327-328: “We nevertheless… gridcells”: can be deleted as it has already been said in the method part.   
- Fig. 4: This figure could be larger. “Topographic parameters” could be added for the x-axis.   
- Fig. 5: Information about the date (winter 2016, measurements performed in may) could be added in the legend.   
In this figure the LR method used remains unclear. Have topographic parameters been used?  I also don’t really understand how this figure indicates that elevation and Sx are the most significant predictors (mentioned in line 333).   
- Lines 331-335: We can also see that the larger correlation is obtained for the glacier covering the largest elevation ranges (and vice versa). This could be mentioned. In addition melting events don't seem to affect the glacier 4 as stronger as for the two other one (also mention below).   
So do you think that the low correlation with the elevation for glacier 4 is due to the absence of snow melt event or that the glacier covers low elevation ranges, or a combination of both? To provide better precision, the correlation could be computed in the upper part of the glacier less affected by melting events. I think it is important and relevant to discuss this point, as most of the studies indicate the elevation as the main predictor of the WB variability, but it could be principally due to winter melting events...   
- Line 335-342: This sentence provides general information and should be mentioned in the introduction. References to other studies should be mentioned as a comparison with your results/ to discuss your results (e.g. “Our results are in good agreement with previous studies (references) which have found the elevation to be the…”   
Same remark regarding the variability of WB-gradients between glaciers: your findings should be mentioned here to be compared with the reference mentioned.   
Same remark regarding slope, wind (mentioned then at lines 346-361), ect…   
I think you really have interesting results, but it is quite hard to follow the discussion. I therefore suggest reorganizing this paragraph (from line 335 to 370) and splitting it into a result and a discussion section (as mention in my main comment). Here, only mention the main results of linear regression should be included. In the discussion part, discuss the spatial variability of WB (at glacier scale / between glaciers / at regional scale) and compare your result with previous studies. Note that the section “Regional winter balance gradient” as well as Fig.9 could be insert in this section.   
Lines 360-361: Yes, but how? Using models?   
Line 380: The kriging method always gives very bad extrapolations!   
Line 386-387: I understand that the 1/3 kept was only necessary to compute LR uncertainty, so the LR method is based on 2/3 of the data. So to make the comparison more consistent, the KG uncertainty could also be computed using only 2/3 of the data.   
Line 393: Probably because there is less data measurements in the accumulation area. As KG extrapolations usually lead to large uncertainties, maybe the comparison between the 2 approaches could be reduced over the measurement area (i.e. only based on interpolations).   
Line 397: interpolation uncertainties or interpolations and extrapolations?   
Line 397-398: So why not computing topographic regression only over the measurement area?   
Figure 8: Some colors are really difficult to see. Please use more contrasted colors.   
Line 405: Same remark than above: the comparison could be only based on the measurement area using the interpolation (and not the extrapolation).   
Table 4: In the legend please add (WB) to define the acronym.   
Line 415: “…from a number of sources” Please indicate what are these sources (e.g. snow depth measurements?)   
Line 418: It is unclear according me. Could you provide more information?   
Line 422: What are here the “combined sources of uncertainty”?   
Line 425-437: I would recommend to include this paragraph in a section discussing the spatial variability of WB (please refer to my remark above and my main comment.   
Lines 454-455: This could be mentioned in introduction.   
  
Conclusion   
Line 507: I suggest to mention the principal limit of the study as well as future works here.   
  
Supplementary material   
General comment: The method is here briefly described. As you are not limited, you could provide more information.   
More detailed comments:   
- The method to obtain the smoothed DEM is unclear. Could you provide more information?   
- Can you provide more information about the Sx index computation? For instance, is different angle sizes and distances have been tried?   
  
  
Reviewer: 2   
  
Comments to the Author   
see attachment

Report on ”Estimating winter balance and its uncertainty from direct measurements of snow depth and density on alpine glaciers”

Alexandra Pulwicki, Gwenn E. Flowers, Valentina Radic

The aim of this work is to estimate accurately winter surface mass balance and evaluate its uncertainty on three glaciers in the St. Elias Mountains, Yukon, Canada by statistical methods and from direct measurements. 9000 manual measurements have been collected through a sophisticated sampling design aiming at catching both the variability at different scales and the heterogeneity across the glaciers. The paper is well written and well organized, the figures and tables are relevant and well documented.

*Thank you!*

I am not specialist in glaciers study, rather in environmental statistics, I will not comment extensively the part of the paper concerning the measurements and the discussion on the interpretation of the results, just to say that as far as I can judge the goal of the study is an important point to assess, the measurements collected form a very valuable dataset for this domain of research and that this paper and its results deserve being published. I will focus mainly on the statistical methods, for which I have several concerns.

1. First it is likely obvious for those who work on glaciers, but a definition of glacier-wide winter balance and how it is derived from the distributed estimates of winter balance is lacking. I guess it is the average of the distributed values, but it should be precised.

*This is a good observation. We have added the following sentence in the first paragraph of the Methods section to specify that we took the average of all gridcells.*

“Glacier-wide WB is then calculated by taking the average of all gridcell-averaged WB values for each glacier.”

*Further, we have decided to use to indicate glacier-wide winter balance throughout the manuscript to be consistent with the Glossary of glacier mass balance and related terms (Cogley and others, 2011). This change in notion was informally suggested by colleagues after submission. We hope that using will improve clarity of the process used to calculate glacier-wide winter balance. Point- and gridcell-scale winter balance are now denoted as to further comply with the Glossary of glacier mass balance and related terms.*

*Cogley J, Hock R, Rasmussen L, Arendt A, Bauder A, Braithwaite R, Jansson P, Kaser G, Moller M, Nicholson L and others (2011) Glossary of glacier mass balance and related terms. UNESCO-IHP, Paris*

1. I did not find an information that is crucial for me: in how much gridcells were there point-scale values to average and what is the total size of the 40 x 40m DEM grid? It is stated line 183 that the point-scale values are averaged in each 40 x 40m DEM gridcell, but I suppose that there is not point-scale values in each gridcell.

*We agree that this is important information to include. We have added the requested information to Table xx.*

1. About the Linear Regression procedure, I am not sure on how it is really performed. It is stated that first, cross validation is used to obtain a set of βi and second model averaging is used to account for uncertainty when selecting the predictors, but how are estimated the βi in the cross validation step? If it is by model averaging, then the two step are nested and the presentation of the procedure needs to be reformulated.
2. I have more concerns with the simple kriging part.

(a) Simple kriging needs the knowledge of the expectation of the random field in all the sites to be predicted and which are measured. How is it estimated? If it is assumed constant over the glacier its value should be (in my opinion) the glacier- wide winter balance and this is just what is intended to be estimated from the kriged values.

(b) No model is given for the covariance function, as the fitted parameters are the range and the nugget, I guess that it would be an exponential or a Gaussian covariance function, but this needs to be clarified and justified.

*We have added details about the covariance function.*

“A Matere covariance function with ν=5/2 is used to define a stationary and isotropic covariance and covariance kernels are parameterized as in Rasmussen and Williams (2006).”

(c) For that kind of data it is usual to use rather ordinary kriging with a fitted variogram, the advantage is that the expectation can be unknown and needs not being estimated, and the stationarity assumption is relaxed to an assumption of stationary increments.

(d) Even more sophisticated kriging may be performed accounting for a trend, for in- stance universal kriging that use the geographical coordinates, or regression kriging which kriges the residuals of a regression, it could be for instance the LR previously introduced.

5. The uncertainty analysis is driven by different means according to the steps of the procedure. It is indicated at the beginning of the section that repeated random sampling is performed according to different input variables. It is the case for the uncertainty linked to the grid scale variability, 1000 values are drawn from a zero mean normal distribution with standard deviation equal to the mean standard distribution of zigzag.

(a) The inputs to evaluate the density assignment uncertainty are the eight density interpolation methods hence eight values, is then the density assignment uncertainty the standard deviation of the eight resulting values for the glacier wide winter balance? This would be very far from the announced 1000 sample values.

*This comment is correct, we use eight values to calculate the standard deviation for density assignment uncertainty. To clarify this confusion, we have rearranged the information presented in the first paragraph of the Methods: Uncertainty analysis section and added the following sentence:*

“Density assignment uncertainty is calculated as the standard deviation of the eight resulting values of glacier-wide winter balance.“

(b) The way that is calculated the interpolation uncertainty for the simple kriging is weird. A standard deviation is derived from the confidence interval given by the package DiceKriging, but usually the confidence interval is calculated from the kriging standard deviation, why this value is not used directly? I don’t know if it is an output for this package, but it is without any doubt for most standard kriging packages (gstat, geoR, spatial, fields, RandomFields ...).

*Upon investigation, DiceKriging does directly output the standard deviation of estimated values. The results for simple kriging are equivalent when the standard deviation is used directly and when the standard deviation is derived from the confidence interval. The section that details the calculation of simple kriging uncertainty now reads as follows:*

“SK interpolation uncertainty is represented by the standard deviation for each gridcell-estimated value of WB generated by the DiceKriging package. “

The most serious concern on this work, is the use of simple kriging, which seems not relevant in this framework. Some points need to be clarified (number of gridcell averaged values, steps of the LR procedure, calculation of the uncertainty of the density assignment and the interpolation assignment).

I recommend a major revision, as this work develop many interesting points and deserves being published.