We thank the reviewer very much for their feedback on the manuscript. These are the main changes we made in response to your comments; Added a validation of energy balance components to show the reason we scaled the wind speed was to increase the sensible heat flux. Added a discussion of the key strengths and weaknesses of the model (including refence to previous work to implement a glacier scheme into the REMO model). Added a discussion on the drawbacks of our calibration approach. Added a section with figures on mass balance components to show how these vary with height and how these may change in the future. Included percentage volume change projections for elevation levels. Added a discussion on the reasons for the model bias in the seasonal mass balance validation (also commented on by reviewer #1). Please see the detailed replies to your comments below.

Reviewer comment: The proposed manuscript presents new estimates of future sea-level rise obtained with a new model of global glacier evolution. The model is quite original because it is integrated in a land surface model. Also, the mass-balance is computed from the surface energy balance, which could lead to different and more contrasted results than classical "temperature-index" mass-balance models. For these reasons, it is a useful addition to the literature. That said, I think that the paper needs serious revisions before being considered for publication. I have some major concerns listed below.

My recommendation would be to:

- **use additional datasets for model validation**
- **put less focus on the validation with point mass-balance data**
- **spend more time on the energy balance instead**

Reply: To focus more on the energy balance and to justify why we give a high prominence to adjusting wind speed (your point below) we have added a comparison of the modelled energy balance components with observations from the Pasterze glacier in the Alps (Greuell and Smeets 2001). We found that the model underestimated the sensible heat flux by an order of magnitude and the wind speed was four times lower than the observations. The sensible heat flux is underestimated because the surface exchange coefficient (used to calculate the turbulent heat flux) is proportional to the wind speed.

You make the point (further down) that the wind speed scaling does not seem physical because the katabatic wind is decoupled from synoptic conditions. We agree that our approach is rather crude and could certainly be improved. However, we thought it was better to include a crude adjustment for wind speed rather than neglecting this process altogether.

Change to manuscript: We have added the following to model description of wind speed section 2.2.4

"Although our approach is rather crude, we found that scaling the wind speed was necessary to get reasonable values for the sensible heat flux. This is seen when we compare the modelled energy balance components to observations from the Pasterze glacier in the Alps (Greuell and Smeets 2001). The measurements consist of incoming and outgoing short and long wave radiation, albedo, temperature, wind speed and roughness length at five heights between 2205m-3325m meters above sea level on the glacier. Table S6 in the Supplementary Material lists the observed and modelled energy balance components and meteorological data, for experiments with and without wind speed scaling. The comparison shows that JULES underestimates the sensible heat flux by at least one order of magnitude and the modelled wind speed is four times lower than the observations. When we increase the wind speed to match the observations there is a better agreement with the observed sensible heat

flux. This is because the surface exchange coefficient, which is used to calculate the sensible heat flux, is a function of the wind speed in the model. "

Reviewer comment: Model set-up I found it difficult to understand some aspects of the model set-up, and I believe the text could be meliorated by reorganizing some of its sections.

In particular:

• the description of the snowpack/ice meld model is scattered around Section 2/ Intro, and page 7.

Reply: We preferred to keep the description of the snowpack initialisation (page 7) separate to the snowpack description (Section 2). We also kept the description of the multi-level snowpack scheme brief because it is described in detail in Best et al 2011.

Reviewer comment: the description of how and where glaciers lose mass is unclear to me (and to the other reviewer as well)

Reply: We have added the following text early in the model description section to help clarify how and where glaciers loose mass

Change to manuscript: "Each elevated glacier tile has a snowpack which can gain mass through accumulation and freezing of water and lose mass through sublimation and melting. JULES has a full energy balance multi-level snowpack scheme which splits the snowpack into layers each having a thickness, temperature, density, grain size (used to determine albedo), and solid ice and liquid water contents. The initialisation of the snowpack properties and the distribution of the glacier tiles as a function of height is described in section 2.3"

Reviewer comment: The manuscript would also benefit from a clear discussion and listing of the particularities (strengths/weaknesses) of this model in comparison to previous modelling attempts.

This could be a way to highlight the strengths of your model (see "Energy balance" below).

Reply: We modified the discussion section to list the strengths and weaknesses of the model.

Change to manuscript: "There are three key strengths to the JULES glacier model. Firstly, we include variations in orography within a climate gridbox which is important to calculate elevation-dependent glacier mass balance. Kotlarski et al (2010) developed a glacier scheme for the REMO regional climate model by lumping glaciers into 0-5-degree gridboxes in a similar approach to us, but they did not have a representation of subgrid orography. Instead glacier gridboxes received double the gridbox mean snowfall, glacier ice had a fixed albedo and a constant lapse rate was applied to adjust temperatures. They concluded that to reproduce mass balance trends over the Alps, the scheme needed to include subgrid variability of atmospheric parameters within a gridbox.

Secondly, the model uses a full energy balance scheme to calculate glacier melting. This is a more physically based approach than the widely used temperature index models, which relate melting to temperature using a degree day factor (DDF). The DDF lumps all the energy balance components into a single number meaning that the effects of changing wind speed, cloudiness and radiation on melt rates cannot be considered. Changes in solar radiation can be an important driver of melting. Huss et al (2009) studied long term mass balance trends for a site in the Alps and showed that melting was stronger during the 1940's than in recent years despite more warming. This was because summer solar radiation was higher during the 1940s. Moreover, temperature index models have been found to be less accurate with increasing temporal resolution (for example on daily time steps) (Hock 2005).

Finally, the glacier scheme is coupled to a land surface model, which presents opportunities for further studies. For instance, the model could be used to investigate the impact of climate change on river discharge in glaciated catchments in Asia, South America or the Arctic."

One of the major shortcomings of the model is that glacier dynamics is not included (glacier area does not vary). The model does not simulate the retreat of the glacier terminus which results in an overestimate of mass loss. Neither does the model simulate the transport of ice from higher elevations to lower elevations.

An additional drawback of the model is the coarse resolution of the gridboxes which make it unfeasible to include some process which affect local mass balance such as hillside shading, avalanching, blowing snow and calving. The model could, however, be run on a finer resolution using higher resolution climate forcing data."

Reviewer comment: Calibration/validation

Validating and calibrating against in-situ elevation dependant MB data is really hard. Point mass-balance observations reflect a number of local and glacier specific factors, and there is a risk that these local factors affect parameter sets chosen for entire RGI C2 regions (this is particularly true for regions with few observations). I am not asking to change your procedure as this will likely represent too much work, but I'm allowing myself to provide a couple of suggestions:

I would personally recommend to use one global set of parameters instead of the regional ones, unless there is a compelling reason not to do so (a good candidate would be a physical explanation of the regional parameter sets). The resulting parameter sets would be more robust and would allow statistical scrutinizing using cross-validation (or more advanced) methods. This is even more relevant for physically based approaches like yours.

Reply: It is not clear why a single global parameter set would be more robust than regional parameters sets. For example, in Table 2 we show that the optimal value for the fresh snow albedo in the visible range is 0.83 in Central Europe and 0.97 in Western Canada and the US. A single parameter set would not capture this regional variation. Similarly, the wind speed scaling varies between the regions, 1.83 in Europe and 2.29 in Western Canada and the US.

Reviewer comment: Having some kind of independent validation would considerably strengthen the readers' confidence in your results. Albeit not without their own problems, regional geodetic mass-balance estimates could be useful to at least get a quantitative estimate of the model performance at the regional scale.

Reply: We agree with the reviewer that the calibration approach could be improved so we have added a section on this in the discussion.

Change to manuscript: "The robustness of the glacier projections depends on how well the model can reproduce present-day glacier mass balance. One of the main shortcomings of the calibration and validation of mass balance is that only a single type of observations is used. This data was used because we wanted to ensure the model could reproduce variations in accumulation and ablation with height when the elevated tiling scheme was introduced. Point mass balance observations are affected by local factors such as aspect, avalanching, debris cover and there is a possibility that these local factors affect *parameter sets chosen for entire RGI region. This could be improved by using observations from satellite gravimetry and altimetry, such as that described by Gardner et al (2013) to get a quantitative estimate of the model performance at the regional scales. "*

Reviewer comment: Energy balance

The real strength of the model is its use of an energy balance model instead of a temperature index model as the majority of the other global models. Whether or not this increase in complexity is actually leading to better results remains (and will remain) a controversial topic, but this study should make use of this novel approach. In particular, I would find it very interesting to see plots of the energy balance components as a function of altitude, and how these energy balance components change in the future. This is interesting because energy balance models are likely to be less sensitive to temperature change and incorporate other processes instead. I would also welcome new analyses of not only the total volume change, but the volume changes per elevation band.

Reply: This extra analysis is added to the paper

Change to manuscript: We added percentage volume changes for low, medium and high elevation ranges to Table 6 and added the following text to section 4.2

"The percentage volume changes for three different elevation ranges; low (0-2000m), medium (2250m-4000m) and high (4250m-9000m) are listed in Table 6. Some of the high latitude regions particularly Alaska, Western Canada & US, Svalbard and North Asia experience very large volume increases at their upper elevation ranges. This would be reduced if the model included glacier dynamics, because ice would be transported from higher elevations to lower elevations."

Change to manuscript:

"4.3 Mass Balance Components

In this section we examine how the surface mass balance components vary with height and how this will change in the future. Fig. 12 shows the accumulation, refreezing and melting contributions to mass balance averaged over low, medium and high elevations ranges for the period 1980-2000. Sublimation is excluded because its contribution to mass balance is relatively small. As expected there is more melting in the lower elevation ranges and more accumulation at the higher elevation ranges. The refreezing component, which includes refreezing of melt water and elevated adjusted rainfall, shows no clear variation with height. This is because the refreezing component can both increase and decrease with height. Refreezing can increases towards lever elevations because there is more rain and melted water. It can also decrease if the snowpack is depleted or if there is not enough pore space to hold water because previous refreezing episodes have converted the firn into solid ice. The largest accumulation rates occur in Alaska (5.3 m.w.eq.yr-1) and Western Canada and US (7.3 m.w.eq.yr-1) between 4250m-9000m and the largest melt rates are found in the Caucasus and Middle East (-7.4 m.w.eq.yr-1) and the Low Latitudes (-7.6 m.w.eq.yr-1).

Fig. 13 shows how the global annual mass balance components vary with time for low, medium and high elevations ranges. At the high and medium elevations accumulation, refreezing and melting decrease leading to a reduction in mass loss as glaciers disappear towards the end of the century. At high elevations mass balance is reduced from -2.2 m.w.eq.yr^{-1} (-177 Gtyr-1) during the historical period (1980-2000) to -0.35 m.w.eq. yr-1 (-28 Gtyr^{-1}) by the end of the century (2080-2097). Similarly, for *the medium elevation ranges mass balance reduces from -0.56 m.w.eq.yr^{-1} (-26 Gtyr^{-1}) to -0.24 m.w.eq.yr^{-1} (-11 Gtyr^{-1})."*

Reviewer comment: Code availability Please add a statement about where and how people can access your code and that of JULES.

Reply: We added a section on code availability at the end of the paper.

Change to manuscript:

Code availability

The glacier scheme is included in JULES v4.7. The source code can be downloaded by accessing the Met Office Science Repository Service (MOSRS) (requires registration): https://code.metoffice.gov.uk/ The code used for this study is in https://code.metoffice.gov.uk/svn/jules/main/branches/dev/sarahshannon/vn4.7_va_scaling

Reviewer comment: Specific comments

P3 L15 why is this limitation about the partial coverage necessary? In view of the objective of developing a fully coupled model, it would be good to overcome this limitation one day.

Reply: We agree with the reviewer that it would be preferable to be able to mix the elevated tiles and vegetated tile schemes within gridboxes. There are a number of structural difficulties in the JULES code that make it very difficult to do this unfortunately, well outside the scope of even the significant code development that was done for the elevated tiles used in this study.

One of the main difficulties is that the primary soil model in JULES (which is essential for the vegetation, but incompatible with glaciated tiles) is fundamentally structured to run with one set of variables and parameters for each gridbox. Thus, for each gridbox one must choose to have either a soil or an ice subsurface, and all the tiles above for that gridbox must fit exclusively into one of those two categories. Work has been ongoing for a number of years in the JULES development groups to allow the subsurfaces (like soil) to be tiled like the surfaces, but this is a major undertaking which touches almost every part of the codebase and as yet there is no firm timescale for this work to be completed. In addition, the surface tiles in JULES can change area fraction during the course of a run, as climate favours different vegetation types or a glacier changes in volume, but extending this essential functionality to the tiled soils brings non-trivial challenges in carbon and water conservation within the model that have yet to be addressed.

Reviewer comment: P3 L23 0.5◦ and 46 elevation bands: What motivated the choice of these resolutions? Can this be changed at wish?

Reply: Yes, these resolutions can be changed. The 0.5^{o} resolution was chosen because the climate data (historical and future) is on this resolution. The vertical resolution of 250m was chosen for computational cost. This could be increased to a finer vertical resolution, but a new tile fraction ancillary and initial snowpack depth would need to be generated from the 50m glacier hypsometry data in the RGI6.

Change to manuscript: Added text to the model description: "The horizontal resolution of 0.5-degree is used because it matches the forcing data used to drive the model. The vertical resolution of 250m was used based on computational cost. The vertical and horizontal resolutions of the model can be modified for any setup."

Reviewer comment: P4 L5 Snowpack: do I get this right that there is no distinction between ice and snow in the snowpack model? What are typical values for ice density in the model? How much time does it need to transform snow to ice?

Reply: There are some distinctions between snow and ice in the snowpack. Snow and ice have different densities, ice and liquid water content, grain size (which determines albedo) and temperature. Fresh snow at the top of the snowpack has a typical density 250kgm^{-3} and ice has a density 917kgm^{-3}. The albedo is treated differently for ice and snow. The new albedo scheme, which scales the albedo as a function of the snowpack surface density, is activated when the firn density is greater than 550kgm^{-3} Below this threshold the snow aging scheme is used.

We have not estimated how long it would take for snow to convert to ice, because in our model setup we always prescribe the bottom of the snowpack with solid ice. However, theoretically the change in density and grain size with time is described in Best et al 2011 (Equations 21 and 39 respectively). It would be interesting to explore whether the model can grow realistic glaciers from scratch and if so, how long this would take.

Reviewer comment: P4 L10 I must admit that I dislike the current approach to temperature downscaling, which in my opinion is an unhealthy mix of thermodynamics and tuning. I'm not asking to change it, but the fact that the lapse-rate is tuned in non-saturated conditions (the rate changes quite a lot according to table 2) but not in saturated conditions is likely to create odd non-linearities in the model's response to certain forcings. P5 L1 "we only tune the dry adiabatic lapse-rate". Here and throughout the rest of the manuscript: do not use the term "dry adiabatic lapse-rate". The dry adiabatic lapse-rate is the dry adiabatic lapse-rate and is 9.8K per 1000m. What you are tuning though is the near-surface temperature lapse-rate, which might vary according to surface conditions and moisture content.

Change to manuscript: We removed "P5 L1 "we only tune the dry adiabatic lapse-rate" and replaced instances of "dry adiabatic lapse-rate" with "lapse rate"

Reviewer comment: P5 L22 wet-bulb temperature is a much better indicator for solid precipitation than regular dry-bulb temperature. This could mitigate parts of the dramatic changes in snowfall projected by your scenarios.

Reply: We agree with the reviewer that using wet-bulb temperature to partition rain and snow is better than dry-bulb temperature. We did preliminary sensitivities studies which showed that the mass balance was more sensitive to the precipitation lapse rate than the temperature for partitioning rain and snow. However, one of the reasons we have a negative bias in the seasonal mass balance (an overestimate in melting and an underestimate in accumulation) is because we use a dry bulb temperature of ^{0}C which is likely too low for partitioning rain and snow. We have added the following text to the model validation section 3.2

Change to manuscript: "Some, but not all, of the bias is due to the partitioning of rain and snow based on an air temperature threshold of 0^{o}C. The 0^{o}C threshold is likely too low, resulting in an underestimate of snowfall. When precipitation falls as rain or snow it adds liquid water or ice to the snowpack. The specific heat capacity of the snowpack is a function of the liquid water (W_k) and ice content (I_k) in each layer

$$\boldsymbol{C_k = I_k C_{ice} + W_k C_{water}} \qquad (17)$$

where C_{ice} = 2100 $JK^{-1}kg^{-1}$ and C_{water} = 4100 $JK^{-1}kg^{-1}$. The liquid water content is limited by the available pore space in the snowpack, therefore changes in the ice content control the overall heat capacity. The underestimate in the ice content reduces the heat capacity which causes more melting than observed.

Other modelling studies have used higher air temperature thresholds; 1.5°C (Huss and Hock 2015, Giesen and Oerlemans 2012), 2°C (Hirabayashi et al 2010) and 3°C (Marzeion et al 2012). An improved approach would use the wet-bulb temperature to partition rain and snow which would include the effects of humidity on temperature. Alternatively, a spatially varying threshold based on precipitation observations could be used. Jennings et al (2018) showed by analysing precipitation observations, that the temperature threshold varies spatially and generally higher for continental climates than maritime climates.

Reviewer comment: P5 L27 here and at some other places in the manuscript, the missing katabatic flow is given a high prominence in the list of missing processes to be addressed. This might be the case (also not the most prominent on my list), but the proposed solution (scaling the synoptic wind field) does not sound really physical to me. The katabatic flow is notoriously decoupled from the synoptic conditions and is likely to be strongest when the synoptic flow is weak. What the scaling of the modelled wind achieves, though, is an increase of the turbulent fluxes: I would be very interested in seeing more discussion about why this is necessary (see major point 1.3 above).

Reply: The reason we include katabatic winds was to increase the modelled sensible heat flux compared to observations. Please see our first point (above) responding to this.

Reviewer comment: P7 L1-2 If I get this right, the glacier tiles are able to lose mass per elevation band, right? The information is scattered in the manuscript (P7 L10, P7 L24 . . .) and should be clarified much earlier to avoid confusion (see also the comment from reviewer 1 who seems to have understood something different than me).

Reply: Yes, you understand correct.

Change to manuscript: We add this line early in the model description at page 3 to help clarify "This allows glacier tiles to gain or lose mass at elevation bands"

Reviewer comment: P7 L10 I don't understand this part. Can you be more specific about how glaciers grow/shrink and lose/gain mass in the model?

Reply: In this line we refer to the glacier tile fraction which does not grow/shrink because the glacier area is fixed. The tile fraction (i.e. the fraction of a gridbox covered in ice at elevation levels) is calculated from the glacier area. Glaciers gain/lose mass though melting, sublimation, accumulation and refreezing in the snowpack but the area (tile fraction) remains fixed.

Reviewer comment: P7 L25 What happens at the end of the initialisation? Setting 500 m of ice everywhere is maybe ok for a spin-up, but what happens next, or at the start of the 2011-2100 simulation?

Reply: For the future simulations the depth of the bottom level of the snowpack comes from the RGI6 thickness which is based on the thickness inversion Huss and Farinotti (2012).

Change to manuscript: Added this line to the section 2.3.2 "

For the future simulations the thickness and ice mass at the bottom of the snowpack comes from thickness and volume data in the RGI6. The data is based on thickness inversion calculations from Huss and Farinotti (2012) for individual glaciers which are consolidated onto 0.5-degree gridboxes.

Reply: We noticed a typo in the model description which we have corrected. For the calibration period the initial ice thickness is 1000m not 500m. The spin-up period is 10 years not 1 year. Our description of the spin-up/initialisation is not very clear so we have modified it.

Change to manuscript: "The snowpack temperature profile is calculated by spinning the model up for 10 years for the calibration period and 1 year for the future simulations. The temperature at the top layer of the snowpack is set to the January mean temperature and the bottom layer and subsurface temperature is set to the annual mean temperature. For the calibration period the monthly and annual temperature comes from the last year of the spin-up. Setting the snowpack temperature this way gives a profile of warming towards the bottom of the snowpack representative of geothermal warming from the underlying soil. The initial temperature of the bedrock before the spin up is set to 0°C but this adjusts to the climate when the model spins up. We use these prescribed snowpack properties as the initial state for the calibration and future runs. "

Reviewer comment: P9 -L27 I don't understand the statement "with the notable exception of the low latitude and Central European regions where melting is over estimated". According to Table 3 and the BIAS measure, melt is over-estimated in 9 regions with Svalbard, Southern Andes and New Zealand striking out with more than 1 m negative bias.

C5 Central Europe even has a positive bias. A quick look through the table indicates a general negative bias.

Reply: The reviewer is correct

Change to manuscript: We removed this line and added "Nine out of the sixteen regions have a negative bias in the annual mass balance. Notably Svalbard, Southern Andes and New Zealand underestimate mass balance by 1 m.w.eq.yr^{-1}. "

Change to manuscript: Also, we noticed that the column containing the number of observations did not match those labelled in Figure 3 so we have corrected this.

Reviewer comment: P10 L3 and Figure 4: the explanation about the Maladeta is irrelevant. Figure 4 shows that there are other pink dots around the Maladeta starts, and there is no need for a case by case explanation here.

Reply: There was a mistake in this figure, so we have updated it. The new figure shows our point that the model performs particularly badly for the Maladeta glacier where melting is overestimated.

Change to manuscript: Figure 4 changed.

Reviewer comment: Section 3.2 This does not represent an independent validation because the same data was used for calibration also. What is striking in Table 4 is that all regions now have a significant negative bias both in winter and summer. How can this be explained?

Reply: The fact we are using similar data to validate the model is certainly a weakness of the calibration and we added this point to the discussion (see above). The negative bias (model overestimates melting in the summer and underestimates accumulation in the winter) was also noted by Reviewer #1 and is likely caused by not bias correcting the gridbox mean precipitation before applying the lapse rate adjustment. This causes an underestimate in the snowfall. (Please see our response to reviewer #1 about this).

Reviewer comment: P10 L21 I would like to see more explanations about the "downscaling" procedure. If only SST and sea-ice are used, this sounds a lot more like a full atmosphere GCM simulation to me than a "downscaling" of a GCM product. In particular, what happens to the land-surface components in HadGEM3? What is actually left from the original GCM signal after "downscaling"?

Reply: The reference to downscaling refers to the use of a higher resolution atmosphere model (HadGEM3 GA6.0) to produce new projections consistent with the CMIP5 SST and sea ice projections used to drive these simulations. A more typical usage of the term downscaling (in dynamical terms) might involve the running of a higher resolution limited area regional climate model with boundary conditions supplied by a GCM.

HadGEM3 benefits from an increased horizontal and vertical resolution over the CMIP5 HadGEM2-ES model, and also has substantial changes to the model dynamics. The version of HadGEM3 used here represents a transition between the CMIP5 and CMIP6 versions of the Hadley Centre climate model. The use of a sub-set of CMIP5 model SST and sea ice as drivers allows an exploration of uncertainties in the regional impacts of climate change, but consistent with the CMIP5 simulations.

Reviewer comment: P11 L2 remove "which is suitable for capturing precipitation variability over complex topography"

Reply: We removed this as suggested.

Section 4.3 Parametric uncertainty analysis is one aspect of parameter uncertainty. A further uncertainty would be revealed by doing data-denial experiments (crossvalidation) and assessing the sensitivity of your calibration procedure to those.

Reply: We agree that data-denial experiments would be valuable to assess the calibration uncertainty. We explored the calibration uncertainty in a slightly different way in response in reviewer #1 comments. We selected best parameters from our Latin Hype Ensemble by minimising the bias and RMSE and maximizing the correlation coefficient. The uncertainty in the global volume loss when the extra performance metrics are used to calibrate present-day mass balance, is approximately double the uncertainty arising from the different climate forcings (Fig. 16, Table 7). This shows the calibration approach has a large impact of the results.

Reviewer comment: Section 4.4 Comparison with other studies I would welcome future studies based on this model to use the same forcing data and same conventions as other global glacier models where possible in order to facilitate model intercomparisons.

Reply: We hope that the model can participate in the Glacier Model Inter-Comparison Project (glacierMIP http://www.climate-cryosphere.org/activities/targeted/glaciermip). This project will compare glacier volume projections for a range of global glacier models each using the same climate forcing and initial ice volumes (RGI6).

Reviewer comment: P13 L25 My understanding is that your study is using the global volume estimates provided by Matthias Huss.

Reply: Yes, that is correct.

C6 Supplementary material I suggest to invert the color scale: red seems more intuitive for mass loss.

Change to manuscript: We reversed the colour scale as suggested.

Best, M. J., M. Pryor, D. B. Clark, G. G. Rooney, R. L. H. Essery, C. B. Ménard, J. M. Edwards, M. A. Hendry, A. Porson, N. Gedney, L. M. Mercado, S. Sitch, E. Blyth, O. Boucher, P. M. Cox, C. S. B. Grimmond & R. J. Harding (2011) The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. *Geosci. Model Dev.,* 4**,** 677-699.

Gardner, A. S., G. Moholdt, J. G. Cogley, B. Wouters, A. Arendt, J. A.Wahr, E. Berthier, R. Hock, W. T. Pfeffer, G. Kaser, S. R. M. Ligtenberg, T. Bolch, M. J. Sharp, J. O. Hagen, M. R. van den Broeke & F. Paul (2013) A Reconciled Estimate of Glacier Contributions to Sea Level Rise: 2003 to 2009. *Science,* 340**,** 852-857.

Greuell, W. & P. Smeets (2001) Variations with elevation in the surface energy balance on the Pasterze (Austria). *Journal of Geophysical Research-Atmospheres,* 106**,** 31717-31727.

Huss, M. & D. Farinotti (2012) Distributed ice thickness and volume of all glaciers around the globe. *Journal of Geophysical Research,* 117**,** F04010.

Rye, C. J., I. C. Willis, N. S. Arnold & J. Kohler (2012) On the need for automated multiobjective optimization and uncertainty estimation of glacier mass balance models. *Journal of Geophysical Research-Earth Surface,* 117.