Response to reviewer #1 Ben Marzeion

We would like to thank the reviewer very much for their feedback. In response to your comments we made the following changes to the manuscript: Added a discussion about the limitation of having a fixed terminus elevation and cited Marzeion et al 2014. A discussion on the negative Nash-Sutcliffe numbers and possible explanation for the model bias in the seasonal bass balance validation. We explored the impact of bias correcting the present-day annual mass balance, on future volume loss projections. We also re-ran the future simulations when the model is calibrated by minimising the bias and RMSE and maximising the correlation coefficient. Please see our detailed relies to your comments below.

Reviewer comment: The model presented here is a very timely and relevant contribution to the growing group of global glacier models since it is, to my knowledge, the only global model that has been developed in the framework of a land surface model. It thus follows a concept that is very different from most other glacier models, adding important diversity and opening a potential path to coupled modelling of glaciers (i.e., allowing interactions between glaciers and ocean/atmosphere, e.g. through the changed freshwater balance), as well as a more integrated perspective on hydrologic impacts of glacier mass

loss.

In principle, I therefore think the authors present a very valuable contribution. However, there are numerous results (particularly concerning calibration and validation) that require much more in-depth analysis and discussion than presented in the manuscript.

This concerns particularly (i) the substantial (and very consistently negative) bias; (ii) modelled negative winter and positive summer mass balances that are currently uncommented; (iii) the lack of a representation of the terminus altitude-mass balance feedback which might also contribute to stronger mass loss projections, and (iv) the Nash-Sutcliffe coefficients that are included in the tables, are mostly negative in the validation, but not at all explained or discussed in the text. If these issues – outlined below in more detail – are appropriately addressed, I would be able to recommend the manuscript for publication.

Specific comments:

Abstract and Introduction: I understand the desire to express the considered CMIP5 projections in terms of temperature above pre-industrial, which has become a quite common measure with the formulation of the goals in the Paris Agreement. However, most readers will be familiar with the RCP scenarios. I was wondering until page 10 whether a “mixed” scenario was used, e.g. selecting CMIP runs based on warming. A quick statement in the abstract and the introduction that RCP8.5 is used will make things a lot clearer.

Changes to manuscript: In the abstract we included: The CMIP5 models use the RCP8.5 climate change scenario and were selected on the criteria of passing 2°C global average warming during this century.
Changes to manuscript: In the introduction we included: "The CMIP5 models use the Representative Concentration Pathways (RCP) RCP8.5 climate change scenario for high greenhouse gas emissions. "

Reviewer comment: P2 L5: In some regions this is true, but since glacier water release is not causally related to demand for water, this sentence should be reformulated.

Reply: We removed the part "when demand for water is high"

Reviewer comment: P2 L7: Most of the authors are native speakers, which makes me doubt myself – but I learned that when used as a compound adjective, sea level should be hyphenated (as in "sea-level rise")?

Changes to manuscript: We have hyphenated instances of the adjective "sea-level rise"

Reviewer comment: P3 L14ff: It would be nice if the authors could comment on how strongly this limitation affects the usability of JULES. I.e., is it realistic that the new glacier scheme would be used in a default setup of JULES, or would the limitations inflicted on the other surface classes be too strong?

Reply: It is feasible to use the glacier scheme in the default configuration of JULES with a caveat. The model is set up so that glacier surfaces cannot share a gridbox with other surface types. This means that running the model with glaciers switched on, should not affect the other surface types. The caveat is this: in order to get sufficient accumulation in the mass balance profile, we had to lapse rate correct the precipitation. This is a standard procedure in mass balance modelling, but the consequence is that gridbox mean precipitation over glacier gridboxes is not conserved. We tested scaling the precipitation, while conserving the gridbox mean i.e. reducing the precipitation near the surface and increasing it at height, but this did not yield enough precipitation to get a good agreement the mass balance profiles. If you wanted to run JULES with the glacier model switched on, then it is worth bearing in mind that water will no longer be conserved. If the model is being used to simulate river discharge in glaciated catchments, then the precipitation lapse rate could be used as a parameter to calibrate the discharge.
Changes to manuscript: We have added this point to the model description section 2.2.3 in case the reader also has the same question.

P4 L5f: If I understand correctly, this implies that the negative feedback between terminus elevation and mass balance is missing, and that the only way for a melting glacier to reach equilibrium with climate is by melting completely (similar to Slangen & van de Wal, 2011, https://doi.org/10.5194/tc-5-673-2011). This is a major limitation that should be discussed in greater depth. E.g., Marzeion et al. (2014, https://doi.org/10.5194/tc-8-59-2014) find that depending on

scenario and accumulated mass loss, this may contribute a few tens of mm SLE (their Figs. 9 to 11). The differences in Tab. S3 could be plausibly explained by this alone.

I also think that the discussion of the lack of a parameterization of ice dynamics (end of Sec. 4.4) is flawed with respect to this feedback.

Reply: Yes, you understand correctly here. The negative feedback between terminus elevation and mass balance is missing and the only way for a melting glacier to reach equilibrium with climate is by melting completely.
Changes to manuscript: We have added the following text to the results section.
"Another explanation why our model predicts more volume loss than Radic et al (2014) and Huss and Hock (2015) is because there is no retreat of the glacier terminus represented in the model. The only way for glaciers to reach equilibrium with climate is by melting completely. A study by Marzeion et al (2014) showed that models predict more mass loss when the terminus elevation is fixed, than when it is allowed to vary. This is because when the terminus retreats, the area available for melting is reduced, leading to less mass loss. Marzeion et al (2014) found that neglecting terminus elevation changes resulted in an extra few tens of mm SLE depending on RCP scenario."
Changes to manuscript: We removed the flawed section regarding the lack of ice dynamics (end of Sec. 4.4)

Reviewer comment: P4 L19ff: Units of the constants are partly missing or wrong.
Changes to manuscript: Units corrected and added.

Reviewer comment: P4 L28: Goff-Gratch (typo), and Landolt-Bordstein 1987 is incomplete in the references, and probably should be Landolt-Börnstein.
Changes to manuscript: Corrected

Reviewer comment: P5 L25: It would be better to say that the energy balance calculated in JULES includes the sensible heat flux, since the snow melt is not a direct (or separate) consequence of the sensible heat flux alone.
Changes to manuscript: Corrected –Instead we say "A component of the energy for melting the snowpack in JULES comes from the sensible heat flux"

Reviewer comment: Eq. 13: I think this equation is a bit problematic if justified by katabatic winds, since the katabatic winds should not be expected to be proportionate to the large-scale wind field.

Reply: We agree. Reviewer #2 also identified this as an unrealistic way to represent katabatic winds (Please see our response to this). We scaled the wind speed to increase the sensible heat flux compared to observations. Although our approach is not a physically realistic, we thought it was better to include a simple way to increase wind speed over glacier gridboxes, than excluding this.

Reviewer comment: P6 L25: "at the beginning" (typo)

P9 L 11: In Marzeion et al. (2012), it is 3 %/100 m.

Changes to manuscript: Corrected

Reviewer comment: Tab. 2: Since the Nash-Sutcliffe efficiency coefficient is included, it should be discussed in the text. It would be particularly good to address the reasons for the numerous negative values – are they caused by bias, too great/small variance, etc.?

Reply: We calculate the Nash-Sutcliffe efficiency coefficient equation using the following equation

$$NS = 1 - \frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{model}\right)}{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{mean\ obs}\right)}$$

where the numerator is the mean square error (or the bias) and the dominator is the variance. The negative numbers arise because the bias is greater than the variance of the observations. The negative bias indicates that melting is overestimated in the summer and accumulation is underestimated during the winter (Table 4).

Changes to manuscript: We added the following text in the model validation section 3.2 to explain the possible reasons for the bias.

"For all regions, except Scandinavia in the summer, negative Nash-Sutcliff numbers are calculated for winter and summer elevation-dependent mass balance (Table 4). The negative numbers arise because the bias in the model is larger than the variance of the observations. There are negative biases for nearly all regions implying that melting is overestimated in the summer and accumulation is underestimated in the winter.

Some, but not all, of the bias is due to the partitioning of rain and snow based on an air temperature threshold of 0ºC. The 0ºC 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 (k)

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

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. Jennings et al (2018) showed by analysing precipitation observations, that the temperature threshold varies spatially and generally higher for continental climates than maritime climates.

Increasing the temperature threshold only reduces the bias slightly, therefore another explanation is that the precipitation in the WFDEI data is too low. Although we have included the variation in precipitation with height, if the gridbox mean precipitation is too low then snowfall on the elevated tiles will be underestimated. We did not bias correct the precipitation before applying the lapse rate correction unlike other studies do (Marzeion et al. 2012, Huss and Hock 2015). The quality of the WFDEI precipitation maybe poor because the data is constrained by rain gauge observations which are sparse in high mountains regions and often biased towards low elevation levels. Even when observations are available snowfall at higher altitudes is often difficult to accurately measure and susceptible to undercatch by 20–50% (Rasmussen et al. 2012). The biases listed in Table 4 are larger in the summer than in the winter. It is likely that the simple albedo scheme, which relates albedo to the density of the snowpack surface, does not perform particularly well in the ablation zone. "

Reviewer comment: Tab. 3: Please add global mean values.

Changes to manuscript: Global values are added. Also, we noticed the number of observations listed in column 5 was incorrect, so we updated this.

Reviewer comment: P10 L1f: The tropical glaciers are really small; there are probably numerous more likely explanations for a warm bias than glaciers lacking in the model.

Figs. 3 and 4, and discussion around them (also P10 L14f): I'm not convinced the Pyrenean glacier is to blame for the low correlation. How much does the correlation change if you exclude it? I'm also wondering why the point cloud for Central Europe in Fig. 3 looks different from the one in Fig. 4?

Reply: There was an error in Figure 4 which we have corrected, and the point cloud now matches the data for Central Europe in Figure 3.

Changes to manuscript: We added the following "In Central Europe some of the poor correlation with observations is caused by the Maladeta glacier in the Pyrenees (Fig. 4) which is a small glacier with an area of 0.52 km^2 WGMS (2017). When this glacier is excluded from the analysis the correlation coefficient increases from 0.26 to 0.35 and the RMSE decreases from 1.99 to 1.73 meters of water equivalent per year."

Reviewer comment: Tab. 4: Please add global mean values.

Changes to manuscript: Values are added

Reviewer comment: Tab. 4: Again, it is necessary to discuss the negative NS-values. There is only one positive value in the table, which indicates that only for summer mass balance in Scandinavia, the model has better predictive skill than taking the mean of the observations. Also, given that all the biases are negative (with the exception of one that is close to zero), the implications for the projections need to be addressed. E.g., if the bias was compensated for in the projections: how would that change the results? Could the differences to previously published projections be explained by the global mean bias?

Reply: Please see above for a discussion on the negative NS-values.

We explored the impact of correcting the bias in the annual mean mass balance on the volume projections. The differences to previously published projections cannot be explained by the bias alone, but it does account for why we have larger volume losses in the Southern Andes, where the bias was particularly large.

Changes to manuscript: The text below has been added to section 4.5 (Comparison with other studies).

"We estimate the end of century global sea-level contribution, excluding Antarctic glaciers, to be 215 ± 20mm which is higher than 188mm (Radic et al. 2014) and 136±23mm (Huss and Hock 2015) caused mainly by greater contributions from Alaska, Southern Andes and the Russian Arctic. These three regions are discussed in turn.

For the Southern Andes our estimates are approximately double (14.4mm) that of the other studies (5.8mm (Huss and Hock 2015), 8.5mm (Radic et al. 2014)). This region has the largest negative bias in the calibrated present-day mass balance (-2.87 m.w.eq.yr^{-1} see Table 3). To explore the effects of correcting the calibration bias on the ice volume projections, we subtract the bias values listed in Table 3 from the future annual mass balance rates. Each gridbox is assumed to have the same regional mass balance bias. The bias corrected volume losses are listed in Table S3 in the supplementary material. For the Southern Andes, the volume losses are much closer to the other studies (7.6mm) when the bias is corrected. The impact is less for the other regions where the biases are smaller. For the Russian Arctic our volume losses are higher than the other studies but that should be interpreted with caution because there were no observations available in this region to get a tuned parameter set (global mean parameters where used instead). In Alaska the bias in annual mass balance is small (0.06 m.w.eq.yr^{-1}) so correcting the bias has little effect on the volume loss projection for this region. Applying the bias correction increases the *global volume loss from 215 ± 20mm to 222.5±20.1mm, therefore the difference between our model and the other studies cannot be explained by the bias in the calibration."*

Reviewer comment: Fig. 5: the RMSE mentioned in the caption is missing in the panels.

Changes to manuscript: The caption has been modified so it no longer says the RMSE error values are shown on the plots. The RMSE values are listed in Table 4.

Reviewer comment: Fig. 6: It is surprising that the model produces many substantially negative winter mass balances (and not quite as many positive summer mass balances). This behaviour should be looked into and discussed in the text.

Reply: This suggests there is a warm bias in the winter but no equivalent cold bias in the summer. We think the presence of a winter warm bias in the model should not necessarily mean there is an equivalent cold bias in the summer.

Reviewer comment: Caption of Fig. 6: "number of glaciers" – isn't that the number of grid boxes?

Reply: This is the number of glaciers for which there are observations.

Reviewer comment:

P10 L23: "sensitivities" (typo)

P11 L12: "Arctic" (typo)

Changes to manuscript: Corrected

Reviewer comment:

Fig S7: It is hard to see anything here; perhaps just leave out the East African and Indonesian glaciers (which is sad for sentimental reasons, but I think they are mostly irrelevant for sea level and water availability).

Changes to manuscript: The figure is modified to only show South and Central America.

Reviewer comment: P11 L16:

Changes to manuscript: Figs. S1 to S7.

Reviewer comment: Sec. 4.3: It is good to see that the model results appear to be robust; on the other hand, this may indicate the negative bias may be hard to overcome. In the calibration, minimizing the RMSE was used for identifying the best parameter set(s). Another way of looking into "parametric" uncertainty (in a wider sense) would be to minimize the bias, or to maximize the correlation or the NS coefficient. These experiments might give valuable insights into the causes of the sometimes problematic model performance, which needs to be better explained.

Reply: We investigated the effect of using multiple performance metrics on the calibrated parameters and glacier volume projections

"Another way to explore the uncertainty in the volume projections caused the calibration procedure, is to use different performance metrics to identify best parameters sets. In addition to using RMSE, we calculate best parameter sets by (1) minimising the absolute value of the bias and (2) maximizing the correlation coefficient. The best regional parameter sets are different depending on the choice of performance metric used (See Tables S2 and S3 in the Supplementary material). For twelve regions, minimising the bias results in higher precipitation lapse rates, than when RMSE values are used to select parameters. This suggests the bias in many regions is caused by underestimating the precipitation lapse rates. As discussed above, this could be due to the fact the gridbox mean WFDEI precipitation was not bias corrected. Glacier volume projections are generated by repeating the simulations using these two additional performance metrics to identify best parameter sets. The uncertainty in the global volume loss when the extra performance metrics are used, is approximately double the uncertainty arising from the different climate forcings (Fig. 1~~1~~6, Table 7). When extra performance metrics are used, the upper bound volume loss increases to 281.1 mm sea-level equivalent by the end of the century."

Changes to manuscript: We changed Figure 16 to show the large spread in the global volume loss when extra performance metrics are used in the calibration.

Extra columns are added to Table 7 to list the regional volume losses when we minimise the bias and RMSE and maximise the correlation coefficient.

Tables are added to the Supplementary Information listing the best parameter selected by minimising the bias (Table S2) and maximining the correlation coefficient (Table S3).

Reviewer comment:

P13 L2: "equivalent" (typo)

P14 L27: Reduction in glacier mass, not necessarily in mass balance.

P15 L18: "periphery" (typo)

Changes to manuscript: Typos corrected

Response to anonymous reviewer #2

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 0.5º 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°C. The 0°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

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

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.

Global glacier volume projections under high-end climate change scenarios

Sarah Shannon[1,2], Robin Smith[3], Andy Wiltshire[4], Tony Payne[2], Matthias Huss[5,6], Richard Betts[1,4], John Caesar[4], Aris Koutroulis[7], Darren Jones[8] and Stephan Harrison[8]

[1]School of Geography, The University of Exeter, The Queen's Drive, Exeter, Devon, EX4 4QJ, UK
[2]Bristol Glaciology Centre, Department of Geographical Science, University Road, University of Bristol, BS8 1SS, UK
[3]NCAS-Climate, Department of Meteorology, University of Reading, Reading, RG6 6BB, UK
[4]Met Office, Fitzroy Road, Exeter, Devon, EX1 3PB, UK
[5]Department of Geosciences, University of Fribourg, Fribourg, Switzerland
[6]Laboratory of Hydraulics, Hydrology and Glaciology, ETH Zurich, Zurich, Switzerland
[7]School of Environmental Engineering, Technical University of Crete, Akrotiri, 73100 Chania, Greece
[8]University of Exeter, Penryn Campus, Treliever Road, Penryn, Cornwall, TR10 9FE, UK

Correspondence to: S. R. Shannon (sarah.shannon@bristol.ac.uk)

Abstract.

The Paris agreement aims to hold global warming to well below 2ºC and to pursue efforts to limit it to 1.5ºC relative to the pre-industrial period. Recent estimates based on population growth and intended carbon emissions from participant countries, suggest global warming may exceed this ambitious target. Here we present glacier volume projections for the end of this century, under a range of high-end climate change scenarios, defined as exceeding +2 ºC global average warming relative to the preindustrial period. Glacier volume is modelled by developing an elevation-dependent mass balance model for the Joint UK Land Environmental Simulator (JULES). To do this, we modify JULES to include glaciated and un-glaciated surfaces that can exist at multiple heights within a single grid-box. Present day mass balance is calibrated by tuning albedo, wind speed, precipitation and temperature lapse rates to obtain the best agreement with observed mass balance profiles. JULES is forced with an ensemble of six Coupled Model Intercomparison Project Phase 5 (CMIP5) models which were downscaled using the high resolution HadGEM3-A atmosphere only global climate model. The CMIP5 models use the RCP8.5 climate change scenario and were selected on the criteria of passing ~~JULES is forced with an ensemble of six Coupled Model Intercomparison Project Phase 5 (CMIP5) models which were downscaled using the high resolution HadGEM3-A atmosphere only global climate model. each~~2ºC global average warming during this century. The ensemble mean volume loss at the end of the century plus/minus one standard deviation is, -64±5% for all glaciers excluding those on the peripheral of the Antarctic ice sheet. The uncertainty in the multi-model mean is rather small and caused by the sensitivity of HadGEM3-A to the boundary conditions supplied by the CMIP5 models. The regions which lose more than 75% of their initial volume by the end of the century are; Alaska, Western Canada and US, Iceland, Scandinavia, Russian Arctic, Central Europe, Caucasus, High Mountain Asia, Low Latitudes, Southern Andes and New Zealand. The ensemble mean ice loss expressed in sea-level equivalent contribution is

215.2 ± 21.3 mm. The largest contributors to sea-~~-~~level rise are Alaska (44.6 ± 1.1mm), Arctic Canada North and South (34.9 ± 3.0mm), Russian Arctic (33.3 ± 4.8mm), Greenland (20.1±4.4), High Mountain Asia (combined Central Asia, South Asia East and West), (18.0 ± 0.8mm), Southern Andes (14.4 ± 0.1mm) and Svalbard (17.0 ± 4.6mm). Including parametric uncertainty in the calibrated mass balance parameters, gives an upper bound global volume loss of 281.1~~247.3~~ mm, sea-level equivalent by the end of the century. Such large ice losses will have inevitable consequences for sea-level rise and for water supply in glacier-fed river systems.

1 Introduction

Glaciers act as natural reservoirs by storing water in the winter and releasing it during dry periods. ~~when demand for water is high.~~ This is particularly vital for seasonal water supply for large river systems in South Asia (Immerzeel 2013, Lutz et al. 2014, Huss and Hock 2018) and Central Asia (Sorg et al. 2012) where glacier melting contributes to streamflow and supplies fresh water to millions of people downstream. Glaciers are also major contributors to sea-~~-~~level rise, despite their mass being much smaller than the Greenland and Antarctic ice sheets (Kaser et al. 2006, Meier et al. 2007, Gardner et al. 2013). Since glaciers are expected to lose mass into the twenty first century (Radic et al. 2014, Giesen and Oerlemans 2013, Slangen et al. 2014, Huss and Hock 2015), there is an urgent need to understand how this will affect seasonal water supply and food security. To study this requires a fully integrated impacts model which includes the linkages and interactions between glacier mass balance, river runoff, irrigation and crop production.

The Joint UK Land Environment Simulator (JULES) (Best et al. 2011) is an appropriate choice for this task because it models these processes, but is currently missing a representation of glacier ice. JULES is the land surface component of the Met Office Global Climate Model (GCM), which is used for operational weather forecasting and climate modelling studies. JULES was originally developed to model vegetation dynamics, snow and soil hydrological processes within the GCM but now has a crop model to simulate crop yield for wheat, soybean, maize and rice (Osborne 2014), an irrigation demand scheme to extract water from ground and river stores and two river routing schemes; Total Runoff Integrating Pathways (Oki 1999)(TRIP) and the RFM kinematic wave model (Bell et al. 2007). The first objective of this study is to add a glacier ice scheme to JULES to contribute to the larger goal of developing of a fully integrated impacts model.

The second objective is to make projections of glacier volume changes under high-end climate change scenarios, defined as exceeding 2 °C global average warming relative to the preindustrial period (2017). The Paris agreement aims to hold global warming to well below 2ºC and to pursue efforts to limit it to 1.5ºC relative to the pre-industrial period (UNFCCC 2015), however, there is some evidence that this target may be exceeded. Revised estimates of population growth suggests there is only a 5% chance of staying below 2 °C and that the likely range of temperature increase will be 2.0–4.9 °C (Raftery et al. 2017). A global temperature increase of 2.6–3.1 °C has been estimated based on the intended carbon emissions submitted by the participant countries for 2020 (Rogelj et al. 2016). Therefore, in this study we make end of the century glacier volume projections, using a subset of downscaled Coupled Model Intercomparison Project phase 5 (CMIP5) models which pass 2ºC and 4°C global average warming. The CMIP5 models use the Representative Concentration Pathways (RCP) RCP8.5 climate change scenario for high greenhouse gas emissions.

The paper is organised as follows; In Section 2 we describe the glacier ice scheme implemented in JULES and the procedure for initialising the model. Section 3 describes how glacier mass balance is calibrated and validated for the present day. Section 4 presents future glacier volume projections, a comparison with other studies and a discussion on parametric uncertainty in the calibration procedure. Section 5 discusses the results, the model limitations and areas for future development. In Section 6, we summarise our findings with some concluding remarks.

2 Model description

JULES (described in detail by Best et al. (2011)) characterises the land surface in terms of sub-grid scale tiles representing natural vegetation, crops, urban, bare soil, lakes and ice. Each grid box is comprised of fractions of these tiles with the total tile fraction summing to 1. The exception to this, is the ice tile which cannot co-exist with other surface types in a grid box. A grid box is either completely covered in ice or not. All tiles can be assigned ~~an~~ elevation offsets from the grid box mean which is typically set to zero as default.

To simulate the mass balance of mountain glaciers more accurately we extend the tiling scheme to flexibly model the surface exchange in different elevation classes in each JULES gridbox. We have added two new surface types, glaciated and unglaciated elevated tiles to JULES (version 4.7) to describe the areal extent and variation in height of glaciers in a gridbox (Fig. 1). Each of these new types, at each elevation, has its own bedrock sub-surface with a fixed heat capacity. These sub-surfaces are impervious to water, and have no carbon content, so have no interaction with the complex hydrology or vegetation found in the rest of JULES. Because glaciated and unglaciated elevated tiles have their own separate bedrock sub-surface they are not allowed to share a gridbox with any other tiles. For instance, gridboxes cannot contain partial coverage of elevated glacier ice and vegetated tiles.

JULES is modified to enable tile heights to be specified in meters above sea level, as opposed to the default option, which is to specify heights as offsets from the gridbox mean. This makes it easier to input glacier hypsometry into the model and to compare the output to observations for particular elevation bands. To implement this change, the gridbox mean elevation associated with the forcing data, is read in as an additional ancillary file. Downscaling of the climate data, described in Section 2.1, is calculated using the difference between the elevation band (z_{band}) and the gridbox mean elevation (z_{gbm})

$$\Delta z = z_{band} - z_{gbm} \quad (1)$$

For the purposes of this study JULES is set up with a spatial resolution of 0.5-degree and 46 elevation bands ranging from 0 - 9000m in increments of 250m. 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 ~~for another type of model setup~~.

Each elevated glacier tile has a snowpack which ~~JULES has a full energy balance multi-level snowpack scheme which~~ can ~~models~~ gain mass through accumulation and freezing of water and lose mass through sublimation and~~,~~ melting.~~,~~ JULES has a ~~compaction and refreezing in the snowpack. The~~ full energy balance multi-level snowpack scheme which ~~is~~ 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. Fresh snow accumulates at the surface of the snowpack at a characteristic low density and compacts towards the bottom of the snowpack under the force of gravity. When rain falls on the snowpack, water is percolated through the layers if the pore space is sufficiently large, while any excess water contributes to the surface runoff. Liquid water below the melting temperature can refreeze. A full energy balance model is used to calculate the energy available for melting. ~~Mass may be removed by sublimation at the surface, or by melting which is implied by the energy balance at each depth.~~ If all the mass in a layer is removed within a model timestep then removal takes place in the layer below. The temperature at each snowpack level is calculated by solving a set of tridiagonal equations for heat transfer with the surface boundary temperature set to the air temperature and the bottom boundary temperature set to the sub-surface temperature.

~~Water can percolate through the snowpack if the pore space is sufficiently large and any liquid water below the melting temperature can refreeze.~~ A ~~S~~snowpack~~s~~ may exist on both glaciated and unglaciated elevated tiles if there is accumulation of snow.

The elevation-dependent mass balance ($SMB_{z,t}$) is calculated as the change in the snowpack mass (S) between successive time steps

$$\boldsymbol{SMB_{z,t} = S_{z,t} - S_{z,t-1}} \quad \textbf{(2)}$$

The scheme assumes that the snowpack can grow or shrink at elevation bands depending on the mass balance, but that tile fraction (derived from the glacier area) is static with time.

2.1 Downscaling of climate forcing on elevations

Both glaciated and unglaciated elevated tiles are assigned heights in meters above sea level and the following adjustments are made to the surface climate in gridboxes where glaciers are present.

2.2.1 Air temperature and specific humidity

Temperature is adjusted for elevation using a dry and moist adiabatic lapse rate depending on the dew point temperature. First the elevated temperature follows the dry adiabat

$$\boldsymbol{T_z = T_0 - \gamma_{dry}\Delta z} \quad \textbf{(3)}$$

where T_0 is the surface temperature, γ_{dry} is the dry adiabatic temperature lapse rate (°Cm^{-1}) and Δz is the height difference between tile elevation and the gridbox mean elevation associated with the forcing data.

If the T_z is less than the dew point temperature T_{dew} then the temperature adjustment follows the moist adiabat. A moist adiabatic lapse rate is calculated using the surface specific humidity from the forcing data

$$\boldsymbol{\gamma_{moist}} = \frac{\left(\frac{g(1+lc.q_0)}{r.T_v(1-q_0)}\right)}{\left(\frac{C_p+lc.2.q_0.R}{r.T_v2.(1-q_0)}\right)} \tag{4}$$

q_0 is the surface specific humidity, l_c is the latent heat of fusion of water at 0°C (2.501 x 10^6 J kg^{-1})~~)~~, g is the acceleration due to gravity (9.8 ms^{-2}), r is the gas constant for dry air (287.05 kg K^{-1})~~,~~ R is the ratio of molecular weights of water and dry air (0.62198) and T_v (K) is the virtual dew point temperature

$$\boldsymbol{T_v = T_{dew}\left(1 + \left(\frac{1}{R} - 1.0\right).q_0\right)} \tag{5}$$

The height at which the air becomes saturated z is

$$\boldsymbol{z = \frac{T_0 - T_{dew}}{\gamma_{dry}}} \tag{6}$$

The elevated temperature following the moist adiabat is then

$$\boldsymbol{T_z = T_{dew} - (\Delta z - z)\gamma_{moist}} \tag{7}$$

Additionally, when $T_z < T_{dew}$, the specific humidity is adjusted for height. The adjustment is made using the elevated air temperature and surface pressure from the forcing data using a lookup table based on Goff-Gratch formula ~~Goff-Grate formula~~ (Landolt-Bornstein 1987). The adjusted humidity is then used in the surface exchange calculation. ~~For simplification we only tune the dry adiabatic lapse rate in this study (see Section 3).~~

2.2.2 Longwave radiation

Downward longwave radiation is adjusted by assuming the atmosphere behaves as a black body using Stefan-Boltzmann's law. The radiative air temperature at the surface $T_{rad,0}$ is calculated using the downward longwave radiation provided by the forcing data $LW_{\downarrow z0}$

$$\boldsymbol{T_{rad,0} = \left(\frac{LW_{\downarrow z0}}{\sigma}\right)^{\frac{1}{4}}} \tag{8}$$

Where σ is the Stefan-Boltzmann constant (5.67 x 10^{-8} W m^{-2} K^{-4}). The radiative temperature at height is then adjusted

$$\boldsymbol{T_{rad,z} = T_{rad,0} + T_z - T_0} \tag{9}$$

Where T_0 is the grid box mean temperature from the forcing data and T_z is the elevated air temperature. This is used to calculate the downward longwave radiation $LW_{\downarrow z}$ at height

$$\boldsymbol{LW_{\downarrow z} = \sigma T_{rad,z}^4} \tag{10}$$

An additional correction is made to ensure that the gridbox mean downward longwave radiation is preserved

$$\boldsymbol{LW_{\downarrow z} = LW_{\downarrow z} - \sum_{z=1}^{n} LW_{\downarrow z}.frac(z)\frac{|LW_{\downarrow z}|}{\sum_{i=1}^{z}|LW_{\downarrow z}|.frac(z)}} \tag{11}$$

where frac is the tile fraction.

2.2.3 Precipitation

To account for orographic precipitation, large scale and convective rainfall and snowfall are adjusted for elevation using an annual mean precipitation gradient (%/100m)

$$P_z = P_0 + P_0\,\gamma_{precip}(z - z_0) \quad \textbf{(12)}$$

where P_0 is the surface precipitation, γ_{precip} is the precipitation gradient and Z_0 is the grid box mean elevation. Rainfall is also converted to snowfall when the elevated air temperature T_z is less the melting temperature (0°C). The adjusted precipitation fields are input into the snowpack scheme and the hydrology subroutine. When calibrating the present-day mass balance, we needed to lapse rate correct the precipitation to get sufficient accumulation in the mass balance compared to observations. The consequence of this, is that the gridbox mean precipitation is no longer conserved. We tested scaling the precipitation, in a way that conserves the gridbox mean by reducing the precipitation near the surface and increasing it at height, but this did not yield enough precipitation to get a good agreement with the mass balance observations. If the model is being used to simulate river discharge in glaciated catchments, then the precipitation lapse rate could be used as a parameter to calibrate the discharge.

2.2.4 Wind speed

A component of the energy available to melt ice, comes from the sensible heat flux which is related to the temperature difference between the surface and the elevation level and the wind speed.

Glaciers often have katabatic (downslope) winds which enhance the sensible heat flux and increase melting (Oerlemans and Grisogono 2002). It is important to represent the effects of katabatic winds on the mass balance when trying to model glacier melt, particularly at lower elevations where the katabatic winds speed is highest.

To explicitly model katabatic winds would require knowledge of the gridbox mean slope at elevation bands, so instead a simple scaling of the surface wind speed is used to represent katabatic winds. Over glaciated grid boxes the wind speed is

$$u_z = u_0\gamma_{wind} \quad \textbf{(13)}$$

where γ_{wind} is a wind speed scale factor and u_0 is the surface wind speed. The simple scaling increases the wind speed relative to the surface forcing data and assumes that the scaling is constant for all heights.

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 ~~We~~ compare~~ed~~ 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~~. T~~ 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.

~~In JULES snow melting is caused by the sensible heat flux which is related to the temperature difference between the surface and the elevation level and the wind speed. Glaciers often have katabatic (downslope) winds which enhance the sensible heat flux and increase melting (Oerlemans and Grisogono 2002). It is important to represent the effects of katabatic winds on the mass balance when trying to model glacier melt, particularly at lower elevations where the katabatic winds speed is highest. To explicitly model katabatic winds would require knowledge of the gridbox mean slope at elevation bands, so instead a simple scaling of the surface wind speed is used to represent katabatic winds. Over glaciated grid boxes the wind speed is~~

$$\cancel{u_z = u_0 \gamma_{wind}}$$

~~(13)~~

~~where γ_{wind} is a wind speed scale factor and u_0 is the surface wind speed. The simple scaling increases the wind speed relative to the surface forcing data and assumes that the scaling is constant for all heights.~~

Formatted: Normal

2.2 Glacier ice albedo scheme

The existing spectral albedo scheme in JULES simulates the darkening of fresh snow as it undergoes the process of aging (Warren and Wiscombe 1980). The growth rate of the grain is an empirically derived function of the snowpack temperature. The snow aging scheme does not reproduce the low albedo values typically observed on glacier ice, therefore a new albedo scheme is ~~it has been modified for~~ used ~~here~~. The new scheme is a density-dependent parameterization which was developed for the implementation in the Surface Mass Balance and Related Sub-surface processes (SOMARS) model (Greuell and Konzelmann 1994). The scheme linearly scales the albedo from the value of fresh snow, to the value of ice, based on the density of the snowpack surface. The new scheme is used when the surface density of the top 10cm of the snowpack ($\rho_{surface}$) is greater than the firn density (550 kgm^{-3}) and the original snow aging scheme is used when ($\rho_{surface}$) is less than the firn density.

$$\alpha_\lambda = \alpha_{\lambda,ice} + (\rho_{surface} - \rho_{ice})\left(\frac{\alpha_{\lambda,snow} - \alpha_{\lambda,ice}}{\rho_{snow} - \rho_{ice}}\right) \quad (14)$$

$\alpha_{\lambda,\,snow}$ is the maximum albedo of fresh snow, $\alpha_{\lambda,\,ice}$ is the albedo of melting ice, ρ_{snow} is the density of fresh snow (250 kgm^{-3}) and ρ_{ice} is the density of ice (917 kgm^{-3}). The albedo scaling is calculated separately in two radiation bands; visible wavelengths $\lambda = 0.3$–$0.7\mu m$ (VIS) and near-infrared wavelengths $\lambda = 0.7$–$5.0\mu m$ (NIR). The parameters, $\alpha_{vis,\,ice}$, $\alpha_{vis\,snow}$, $\alpha_{nir,\,ice}$, $\alpha_{nir,\,snow}$, γ_{temp}, γ_{precip} and γ_{wind} are tuned to obtain the best agreement between simulated and observed surface mass balance profiles for the present-day (see section 3).

2.3 Initialisation

The model requires initial conditions for (1) the snowpack properties (2) glaciated and unglaciated elevated tile fractions within a gridbox. The location of glacier grid points, the initial tile fraction and the present-day ice mass is set using data from the Randolph Glacier Inventory Version 6 (RGI6) (RGI Consortium 2017). This dataset contains information on glacier hypsometry and is intended to capture the state of the world's glaciers a~~ts~~ the beginning of the 21st century. A new feature of the RGI6 is a 0.5-degree gridded glacier volume and area datasets, produced at 50m elevation bands. Volume was constructed for individual glaciers using an inversion technique to estimate ice thickness created using glacier outlines, a digital elevation model and a technique based on the principles of ice flow mechanics (Farinotti et al. 2009, Huss and Farinotti 2012). The area and volume of individual glaciers have been aggregated onto 0.5-degree grid boxes. We bin the 50m area and volume into elevations bands varying from 0m to 9000m in increments of 250m to match the elevation bands prescribed in JULES.

2.3.1 Initial tile fraction

The elevated glaciated fraction is

$$\boldsymbol{frac_{ice(n)} = \frac{RGI\_area(n)}{gridbox\_area(n)}} \quad \textbf{(15)}$$

where *RGI_area* is the area (km^2) at height from the RGI6, *n* is the tile elevation and *gridbox_area* (km^2) is the area of the gridbox. In this configuration of the model, any area that is not glaciated is set to a single unglaciated tile fraction ($frac_{rock}$) with a gridbox mean elevation. It is possible to have an unglaciated tile fraction at every elevation band, but since the glaciated tile fractions does not grow or shrink, we reduce our computation cost by simply putting any unglaciated area into a single tile fraction.

$$\boldsymbol{frac_{rock} = 1 - \sum_{n=1}^{n=nBands} frac_{ice}(n)} \quad \textbf{(16)}$$

nBands = 37 is the number of elevation bands.

2.3.2 Initial snowpack properties

The snowpack is divided into ten levels in which the top nine levels consist of 5m of firn snow with depths [0.05m, 0.1m, 0.15m, 0.2m, 0.25m, 0.5m, 0.75m, 1m, 2m] and the bottom level has a variable depth. For each snowpack level the following properties must be set; density (kgm^{-3}), ice content (kgm^{-2}), liquid water content (kgm^{-2}), grain size (μm) and temperature (ºK). We assume there is no liquid content in the snowpack by setting this to zero. The density at each level is linearly scaled with depth, between the value for fresh snow at the surface (250kgm^{-3}), to the value for ice at the bottom level (917kgm^{-3}). ~~The ice content of the firn layers is calculated by multiplying the scaled density by the depth.~~

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. The ice mass is calculated from the RGI6 volume assuming an ice density 917kgm^{-3}. For the other layers the ice mass is calculated by multiplying the density by the layer thickness which is prescribed above. ~~ice content in the bottom level is set using the RGI6 volume assuming a constant ice density of 917kgm^{-3}. The depth of the bottom level is the ice content divided by the density.~~ For the calibration period, the ice mass at the start of the run (1979) is unknown. In the absence of any information about this, a constant depth of ~~5~~1000m is used which is selected to ensure that the snowpack never completely depletes over the calibration period. This consists of ~~4~~995m of ice at the bottom level of the snowpack and 5m of firn in the layers above. The ice content of the bottom level is the depth (9495m) multiplied by the density of ice.

The snow grain size used to calculate spectral albedo (see section 2.2) is linearly scaled with depth and varies between 50µm at the surface for fresh snow to 2000µm at the base for ice. The snowpack temperature profile is calculated by spinning the model up for ~~one~~10 years for the calibration period and 1 year for the future simulations. ~~. The~~ 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 ~~. This~~ 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 as the model spins up~~when the model is spun up as part of the calibration procedure~~. We use these prescribed snowpack properties as the initial state for the calibration and future runs.

3 Mass balance calibration and validation

3.1 Model calibration

Elevation-dependent mass balance is calibrated for the present-day by tuning seven model parameters and comparing the output to elevation-band specific mass balance observations from the World Glacier Monitoring Service ~~(**WGMS 2017, Global Glacier Change Bulletin No. 2 (2014-2015).**)~~(WGMS, ~~(~~2017). Calibrating mass balance against in-situ observations is a technique which has been used by other glacier modelling studies (Radic and Hock 2011, Giesen and Oerlemans 2013). For the calibration, annual elevation-band mass balance observations are used because there is data available for sixteen of the eighteen RGI6 regions. For validation, winter and summer elevation-band mass balance is used because there is less data available.

The tuneable parameters for mass balance are; visible snow albedo ($\alpha_{vis,snow}$), visible melting ice albedo ($\alpha_{vis,ice}$), near-infrared snow albedo ($\alpha_{nir,snow}$), near-infrared melting ice albedo ($\alpha_{nir,ice}$), orographic precipitation gradient (γ_{precip}), ~~dry adiabatic~~ temperature lapse rate (γ_{temp}) and wind speed scaling factor (γ_{wind}).

Random parameter combinations are selected using Latin Hyper Cube Sampling (McKay, Beckman and Conover 1979) between plausible ranges which have been derived from various sources outlined below. This technique randomly selects parameter values; however, reflectance in the VIS wavelength is always higher than in the NIR. To ensure the random sampling does not select NIR albedo values that are higher or unrealistically close to the VIS albedo values, we calculate the ratio of VIS to NIR albedo using values compiled by compiled by Roesch et al (2002) ~~et al., (2002)~~. The ratio VIS/NIR is calculated as 1.2 so any albedo values that exceed this ratio are excluded from the analysis. This reduces the sample size from 1000 to 198 parameter sets.

In the VIS wavelength the fresh snow albedo is tuned between 0.99 - 0.7 where upper bound value comes from observations of very clean snow with little impurities in the Antarctic (Hudson et al. 2006, 1994). The lower bound represents contaminated fresh snow and comes from taking approximate values from a study based on laboratory experiments of snow, with a large grain size (110 µm) containing 1680 parts per billion of black carbon (Hadley and Kirchstetter 2012). Visible snow albedos of approximately 0.7 have also been observed on glaciers with black carbon and mineral dust contaminants in the Tibetan Plateau (Zhang et al. 2017). In the NIR wavelength the fresh snow albedo is tuned between 0.85 – 0.5 where the upper bound comes from spectral albedo observations made in Antarctica (Reijmer, Bintanja and Greuell 2001). We use a very low minimum albedo for melting ice (0.1) the VIS and NIR wavelengths to capture dirty debris covered ice.

The ~~dry adiabatic~~ temperature lapse rate is tuned between values of 4.0 – 10°C km^{-1} where the upper limit is determined from physically realistic bounds and lower limit is from observations based at glaciers in Alps (Singh 2001). The temperature lapse rate in JULES is constant throughout the year and assumes that temperature always decreases with height.

The wind speed scaling factor γ_{wind} is tuned within the range 1-4 to account for an increase in wind speed with height and for the presence of katabatic winds. The upper bound is estimated using wind observations made along the profile of the Pasterze glacier in the Alps during a field campaign (Greuell and Smeets 2001). Table S64 in the Supplementary Material contains the wind speed observations on the Pasterze glacier. The maximum observed wind speed was 4.6 ms^{-1} (at 2420 meters above sea-level) while the WATCH-ERA Interim dataset (WFDEI) (Weedon et al. 2014) surface wind speed for the same time period was 1.1ms^{-1} indicating a scaling factor of approximately 4.

The orographic precipitation gradient γ_{precip} is tuned between 5-25%/100m. This parameter is poorly constrained by observations therefore a large tuneable range is sampled. Tawde et al. (2016) estimated a precipitation gradient of 19%/100m for 12 glaciers in the Western Himalayas using a combination of remote sensing and in-situ meteorological observations of precipitation. Observations show that the precipitation gradient can be as high as 25%/100m for glaciers in Svalbard (Bruland and Hagen 2002) while glacier-hydrological modelling studies have used much smaller values 4.3%/100m (Sorg et al. 2014) and ~~4.3~~3%/100m (Marzeion et al (2012). The tuneable parameters and their minimum and maximum ranges are listed in Table 1.

The model is forced with daily surface pressure, air temperature, downward longwave and shortwave surface radiation, specific humidity, rainfall, snowfall and wind speed from the WATCH-ERA Interim dataset (WFDEI) (Weedon et al. 2014). To reduce the computation time, only grid points where glacier ice is present are modelled. An ensemble of 198 calibration experiments are run. For each simulation the model is spun up for 10 years and the elevation-dependent mass balance is compared to observations at 149 fields sites over the years 197~~8~~9-2014.

The elevation-dependent mass balance observations come from stake measurements taken every year at different heights along the glaciers. Many of the mass balance observations in the WGMS are supplied without observational dates. In this case, we assume the mass balance year starts on the 1st October to ends on the 30th September with the summer commencing on the 1st May. Dates in the Southern hemisphere are shifted by six months. The observations are grouped according to standardised regions defined by the RGI6 (Fig. 2). The best regional parameter sets are identified by finding the minimum root mean square error between the modelled mass balance and the observations.

Figure 3 shows the modelled mass balance profiles plotted against the observations using the best parameter set for each region. The best regional parameter sets are listed in Table 2 and the root mean square error, correlation coefficient, Nash–Sutcliffe efficiency coefficient and mean bias are listed in Table 3. 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}. ~~The model can capture the accumulation and ablation rates reasonably well for many regions with the notable exception of the low latitude and Central European regions where melting is over estimated.~~

The model performs particularly poorly for the L~~l~~ow L~~l~~atitude region which has a large RMSE (3.02 m.w.eq.yr^{-1}). This region contains relatively small tropical glaciers in Colombia, Peru, Ecuador, Bolivia and Kenya. Marzeion et al (2012) found a poor correlation with observations in the low latitude region when they calibrated their glacier model using CRU data. They attributed that to the fact that sublimation was not included in their model, a process which is important for the mass balance of tropical glaciers. Our mass balance model does include sublimation, so it is possible the WFDEI data over tropical glaciers is too warm. The WFDEI data is based on ERA-interim reanalysis where air temperature has been bias corrected using CRU data. The CRU data comprises of temperature observations which are sparse in regions where tropical glaciers are located. Furthermore, the quality of the WFDEI data will depend on the performance of the underlying ECMWF model. It is possible that the ECMWF model does not include glacier ice in tropical regions. The absence of ice to cool the lower atmosphere would make the grid box mean temperature too warm. ~~In Central Europe the poor correlation with observations is predominantly caused by the Maladeta glacier in the Pyrenees (Fig. 4) which is a small glacier with an area of 0.52 km^2 WGMS (2017).~~ In Central Europe some of the poor correlation with observations is caused by the Maladeta glacier in the Pyrenees (Fig. 4) which is a small glacier with an area of 0.52 km^2 (WGMS, 2017). When this glacier is excluded from the analysis the correlation coefficient increases from 0.26 to 0.35 and the RMSE decreases from 2.03 to 1.73 meters of water equivalent per year.

Formatted: Superscript

Formatted: Not Superscript/ Subscript

3.2 Model validation

The calibrated mass balance is validated against summer and winter elevation-band specific mass balance for each region where data is available (Fig. 5). ~~The model can reproduce the in-situ mass balance observations reasonably well considering the very coarse resolution of JULES in the horizontal (0.5-degree gridded) and vertical (250m increments) directions.~~ For all regions, except Scandinavia in the summer, negative Nash-Sutcliff numbers are calculated for winter and summer elevation-dependent mass balance (Table 4). The negative numbers arise because the bias in the model is larger than the variance of the observations. There are negative biases for nearly all regions implying that melting is overestimated in the summer and accumulation is underestimated in the winter.

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 (*k*)

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

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 snowfall (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^{o}C (Huss and Hock 2015, Giesen and Oerlemans 2012), 2^{o}C (Hirabayashi et al 2010) and 3^{o}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 ~~in~~for continental climates than ~~in~~ maritime climates. ~~. They found that the 50% rain-snow temperature defined as the temperatures above which precipitation is primarily snow, varied between maritime and continental climates.~~

Increasing the temperature threshold reduces the bias slightly, therefore another explanation is that the precipitation in the WFDEI data is too low. Although we have included the variation in precipitation with height, if the gridbox mean precipitation is too low, then snowfall on the elevated tiles will be underestimated. We did not bias correct the precipitation before applying the lapse rate correction ~~un~~like other studies have done~~do~~ (Marzeion et al. 2012, Huss and Hock 2015). The quality of the WFDEI precipitation maybe poor because the data is constrained by rain gauge observations which are sparse in high mountains regions and often biased towards low elevation levels. Even when observations are available snowfall at higher altitudes is often difficult to accurately measure and susceptible to undercatch by 20–50% (Rasmussen et al. 2012).

~~than winter.particularly~~

Winter mass balance is simulated better than summer mass balance, which is seen by the lower root mean square errors for winter in Table 4~~. For all observational sites, the model tends to overestimate the summer melting which can be seen in Fig. 7(A) and by the negative biases in Table 4.~~ . Furthermore, the biases are larger in the summer than in the winter (Table 4). It is likely that the simple albedo scheme, which relates albedo to the density of the snowpack surface, does not perform particularly well in the ablation zone. There is an improvement in the simulated summer mass balance when the glaciated area increases. This is seen by the improved correlation in Fig. 7(D) in which the validation is repeated but only grid boxes with a glaciated area greater than 500km^2 are considered. This indicates that the overestimation in melting is more pronounced for

Formatted: Superscript

Formatted: Superscript

Formatted: Subscript

Formatted: Subscript

Formatted: Subscript

Formatted: Superscript

Formatted: Superscript

Formatted: Subscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

small glaciated areas than regions with a large ice extent. The overestimate in melting in Central Europe is due mainly to the Maladeta glacier as discussed in Section 3.1.

4 Glacier volume projections

4.1 Downscaled climate change projections

Glacier volume projections are made for all regions, excluding Antarctica, for a range of high-end climate change scenarios. This is defined as climate change that exceeds 2ºC and 4ºC global average warming, relative to the pre-industrial period (Gohar et al., 2017). Six models fitting this criterion were selected from the Coupled Model Intercomparison Project Phase 5 (CMIP5). A new set of high resolution projections were generated using the ~~and downscaled using the~~ HadGEM3-A Global Atmosphere (GA) 6.0 model (Walters et al. 2017). ~~–~~ The sea surface temperature and sea-ice concentration boundary conditions for HadGEM3-A are supplied by the CMIP5 models. All models use the RCP8.5 'business as usual scenario' and cover a wide range of climate sensitivities~~itives~~, with some models reaching 2ºC global average warming relative to the pre-industrial period, quickly (IPSL-CM5A-LR) or slowly (GFDL-ESM2M) (Table 5). The models also cover a range of extreme wet or dry climate conditions. This is important to consider for glaciers in the central and eastern Himalaya which accumulate mass during the summer months due to monsoon precipitation (Ageta and Higuchi 1984) and because future monsoon precipitation is highly uncertain in the CMIP5 models (Chen and Zhou 2015).

The HadGEM3-A data are bias corrected using a trend preserving statistical bias method that was developed for the first Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP) (Hempel et al. 2013). This technique uses WATCH forcing data (Weedon et al. 2011) to correct offsets in air pressure, temperature, longwave and shortwave downward surface radiation, rainfall, snowfall and wind speed but not specific humidity. The method adjusts the monthly mean and daily variability in the GCM variables but still preserves the long-term climate signal. ~~The native resolution of HadGEM3-A is N216 (~60km) which is suitable for capturing precipitation variability over complex topography.~~ The ~~data~~ HadGEM3-A was ~~was~~ bi-linearly interpolated from its native resolution of N216 (~60km), onto a 0.5-degree grid, to ~~fit~~match the resolution of the WATCH forcing data which was used for the bias correction. The daily bias corrected surface fields from the HadGEM3-A are used to run JULES offline to calculate future glacier volume changes. The bias correction was only applied to data up until the year 2097, which means the glaciers projections terminate at this year. A flow chart of the experimental set up is shown in Fig. 8. The HadGEM3-A climate data was generated and bias corrected for the High-End cLimate Impact and eXtremes (HELIX) project.

4.2 Regional glacier volume projections 2011-2097

Glaciated areas are divided into 18 regions defined by the RGI6 with no projections made for Antarctic glaciers because the bias correction technique removes the HadGEM3-A data from this region. JULES is run for this century (2011 to 2097) using the best regional parameter sets for mass balance found by the calibration procedure (Table 2). No observations were available to determine the best parameters for Iceland and the Russian Arctic, therefore global mean parameter values are used for these regions. End of the century volume changes (in percent) are found by comparing the volume at end of the run (2097) to the initial volume calculated from the RGI6. Regional volume changes expressed in percent for low (0-2000m), medium (2250m-4000m) and high (4250m-9000m) and all elevation ranges (0-9000m) is listed in Table 6. The total volume loss over all elevation ranges is also listed in ~~in~~ mm of sea-level equivalent ~~are listed~~ in Table 6 and plotted in Fig. 11. Maps of the percentage volume change at the end of the century, relative to the initial volume are contained in the Supplementary Material in Figs. S1-S7~~6~~.

A substantial reduction in glacier volume is projected for all regions (Fig. 9). Global glacier volume is projected to decrease by 64±5% by end of the century, where the value corresponds to the multi-model mean ± one standard deviation. The regions which lose more than 75% of their volume by the end of the century are; Alaska (-89 ± 2%), Western Canada and US (-100 ± 0%), Iceland (-98 ± 3%), Scandinavia (-98 ± 3%), Russian Arctic (-79 ± 10%), Central Europe (-99 ± 0 %), Caucasus (-100 ± 0 %), Central Asia (-80 ± 7%), South Asia West (-98 ± 1%), South Asia East (-95 ± 2 %), Low Latitudes (100 ± 0 %), Southern Andes (-98 ± 1%) and New Zealand (-88±5%). The HadGEM3-A forcing data shows these regions experience a strong warming. In most regions this is combined with a reduction in snowfall relative to the present day, which is drives the mass loss (Fig. 10). Regions most resilient to volume losses are Greenland (-31 ± 5%) and Arctic Canada North (-47 ± 3%). Both these regions lose volume. In the case of Arctic Canada North snowfall increases relative to the present day which helps glaciers to retain their mass. There is a rapid loss of low latitude glaciers which has also been found by other global glaciers models (Marzeion et al. 2012, Huss and Hock 2015). Our model overestimates the melting of these glaciers for the calibration period (Fig. 3), so this result should be treated with a degree of caution.

~~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.

The ensemble mean global sea level equivalent contribution is 215.2 ± 21.3 mm. The largest contributors to sea~~-~~level rise are Alaska (44.6 ± 1.1mm), Arctic Canada North and South (34.9 ± 3.0mm), Russian Arctic (33.3 ± 4.8mm), Greenland (20.1±4.4), High Mountain Asia (combined Central Asia, South Asia East and West), (18.0 ± 0.8mm), Southern Andes (14.4 ± 0.1mm) and Svalbard (17.0 ± 4.6mm). These are the regions which have been observed by the Gravity Recovery and Climate Experiment (GRACE) satellite to have lost the most mass in the recent years (Gardner et al. 2013).

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. Figure 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 ~~As expected melting increases towards~~ in the lower elevation~~s~~ ranges and more ~~more~~ ~~a~~accumulation ~~increases towards~~at the higher elevation ranges. ~~s.~~

The refreezing component, which includes refreezing of melt water and elevated adjusted rainfall, shows no clear variation with height~~discernible pattern~~. – This is because the refreezing component can both increase and~~or~~ decrease ~~at lower elevations~~with height. ~~It~~Refreezing can -increases towards lever elevations ~~The~~because there is more rain and melted water. It can also decreases~~s~~ if the s~~there is no s~~nowpack is depleted or if there is not enough pore space to hold water ~~-refreezing component increases at lower elevation are there is more rainfall and melt water. The refreezing component decrease towards lower elevations because the under warm conditions snowpack might disappear or under cool conditions there there is no 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}) ~~for the~~between ~~-high~~ 4250m-9000m and the ~~elevation range.~~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}).

-Figure 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. ~~There is a reduction in ma~~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~~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}).

4.4 Parametric uncertainty analysis

The standard deviation in the volume losses presented above are relatively small. This is because only a single GCM was used to downscale the CMIP5 data (HadGEM3-A). The uncertainty in the ensemble mean reflects the impact of the different sea-surface temperature and sea-ice concentration boundary conditions, provided by the CMIP5 models, on the HadGEM3-A climate. Other sources of uncertainty in the projections can arise from the calibration procedure, observational error, initial glacier volume and area, and structural uncertainty in the model physics. It is beyond the scope of this paper to investigate all the possible sources of uncertainty on the glacier volume losses. Instead we discuss the impact of parametric uncertainty in the calibration procedure in the following section.

In the calibration procedure the mass balance was tuned to obtain an optimal set of parameters for each RGI6 region, however, there may be other plausible parameter sets that perform equally well (i.e. for which the RMSE between the observations and the model is small). The principle of 'equifinality', in which the end state can be reached by many potential means, is important to explore because some parameters may compensate for each other. For example, the same mass balance could be reached by increasing the wind scale factor which enhances melting or decreasing the precipitation gradient which would reduce accumulation. To identify the experiments that perform equally well, we identify where there is a step change in the gradient

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript of the RMSE for each RGI6 region. A similar approach was used by Stone et al. (2010) to explore the uncertainty in the thickness, volume and areal extent of the present-day Greenland ice sheet from an ensemble of Latin Hypercube experiments. The step change in the RMSE is identified using the changepoint detection algorithm called findchangepts (Rebecca, Fearnhead and Eckley 2012, Lavielle 2005) from the MATLAB signal processing toolbox. The algorithm is run to find where the mean of the top ten experiments (excluding the optimal experiment) changes the most significantly. For each RGI6 region the step changes in the RMSE are shown in Fig. 12.

JULES is re-run for each of the downscaled CMIP5 experiments and for each parameter set that is defined as performing equally well (See Table S1 in the Supplementary Material for a list of the parameters sets that perform equally well). The volume losses expressed in mm of sea-level equivalent are shown in Fig. 13. The effects of the parametric uncertainty on the volume losses varies regionally, with the largest impact found for Central Europe and Greenland. Regional volume loses including parametric uncertainty in the calibration are summarised in Table 7. Including calibration uncertainty in this way ~~Including parametric uncertainty does not change the multi-model mean global volume loss substantially but it does increase the uncertainty range (Fig. 14). This~~ gives an upper bound of 247.3 mm sea-level equivalent volume loss by the end of the century.

Another way to explore the uncertainty in the volume projections caused the calibration procedure, is to use different performance metrics to identify best parameters sets. In addition to using RMSE, we calculate best parameter sets by (1) minimising the absolute value of the bias and (2) maximizing the correlation coefficient. The best regional parameter sets are different depending on the choice of performance metric used (See Tables S2 and S3 in the Supplementary material). For twelve regions, minimising the bias results in higher precipitation lapse rates, than when RMSE values are used to select parameters. This suggests the bias in many regions is caused by underestimating the precipitation lapse rates. As discussed above, this could be due to the fact the gridbox mean WFDEI precipitation was not bias corrected. Glacier volume projections are generated by repeating the simulations using these two additional performance metrics to identify best parameter sets. The uncertainty in the global volume loss when the extra performance metrics are used, is approximately double the uncertainty arising from the different climate forcings ~~Including parametric uncertainty does not change the multi-model mean global volume loss substantially but it does increase the uncertainty range~~ (Fig. 1~~4~~6, Table 7). When extra performance metrics are used, the upper bound volume loss increases to 281.1 mm sea-level equivalent by the end of the century.

4.5 Comparison with other studies

We compare our end of the century volume changes (excluding parametric uncertainty), to two other published studies which used the CMIP5 ensemble under the RCP8.5 climate change scenario (Huss and Hock 2015, Radic et al. 2014). Other studies exist, but these include the volume losses from Antarctic glaciers which makes a direct comparison difficult (Marzeion et al. 2012, Slangen et al. 2014, Giesen and Oerlemans 2013, Hirabayashi et al. 2013). Huss and hock (2015) listed regional percentage volume change and sea-level equivalent values in their study while Radic et al. (2014) listed sea-level equivalent~~eqiuivalent~~ values only (See the comparison Tables S4~~2~~ and S5~~3~~ in the Supplementary material).

Our end of the century percentage volume losses compare reasonably well to Huss and Hock (2015) for Central Europe, Caucasus, South Asia East, Scandinavia, Russian Arctic, Western Canada and US, Arctic Canada North, North Asia, Central Asia, L~~l~~ow ~~l~~Latitudes and New Zealand but are significantly higher in the Southern Andes, Alaska, Iceland and Arctic Canada South. The uncertainty in our percentage volume losses are smaller than Huss and Hock (2015) because we only use a single GCM to downscale the CMIP5 experiments while Huss and Hock (2015) use 14 CMIP5 GCMs.

We estimate the end of century global sea-level contribution, excluding Antarctic glaciers, to be 215 ± 20mm which is higher than 188mm (Radic et al. 2014) and 136±23mm (Huss and Hock 2015) caused mainly by greater contributions from Alaska, Southern Andes and the Russian Arctic. These three regions are discussed in turn.

For the Southern Andes our estimates are approximately double (14.4mm) that of the other studies (5.8mm (Huss and Hock 2015), 8.5mm (Radic et al. 2014)). This region has the largest negative bias in the calibrated present-day mass balance (-2.87 m.w.eq.yr^{-1} see Table 3). To explore the effects of correcting the calibration bias on the ice volume projections, we subtract the bias values listed in Table 3 from the future annual mass balance rates. Each gridbox is assumed to have the same regional mass balance bias. The bias corrected volume losses are listed in Table S5 in the supplementary material. For the Southern Andes, the volume losses are much closer to the other studies (7.6mm) when the bias is corrected. The impact is less for the other regions where the biases are smaller. For the Russian Arctic our volume losses are higher than the other studies but that should be interpreted with caution because there were no observations available in this region to get a tuned parameter set (global mean parameters where used instead). In Alaska the bias in annual mass balance is small (0.06 m.w.eq.yr^{-1}) so correcting the bias has little effect on the volume loss projection for this region.

Applying the bias correction increases the global volume loss from 215 ± 20mm to 223 ± 20 mm, therefore the difference between our model and the other studies cannot be explained by the bias in the calibration.

~~global bias is 0.19 207.8±20.3mm is still higher than the other studies. For Alaska the bias is very small~~

It is likely that our SLE contributions are higher than the other studies because the climate forcing data is different. ~~It is unsurprising that our SLE contributions are higher than the other studies because t~~The HadGEM3-A model uses boundary conditions from a subset of RCP8.5 ~~the~~ CMIP5 models with the highest warming levels. Furthermore, the HadGEM3-A data has a higher resolution (approximately 60km) than the CMIP5 data which was used by the other two studies. This means our model should, in theory, be more accurate at reproducing regional patterns in precipitation and temperature over complex terrain which is important for calculating mass balance.

Another explanation why we predict more volume loss than Radic et al. (2014) and Huss and Hock (2015) is because there is no retreat of the glacier terminus represented in the model. The only way for glaciers to reach equilibrium with climate is by melting completely. A study by Marzeion et al. (2014) showed that models predict more mass loss when the terminus elevation is fixed than when it is allowed to vary. This is because when the terminus is allowed to retreat, there will be less area available to melt. Marzeion et al (2014) found that neglecting terminus elevation changes resulted in an extra few tens of mm SLE depending on RCP scenario.

~~Another difference is that we consolidate individual glaciers into a coarse gridbox and do not model glacier dynamics. In contrast, the other studies model the dynamics of individual glaciers using an elevation change parameterisation (Huss and Hock 2015) or volume-area scaling (Radic et al. 2014). The effect of not including ice flow in our model is that large glaciers will accumulate mass at high elevations because there is no transport of mass to lower elevations. Furthermore, the mass at low elevations will melt too quickly because ice is not replaced by ice flow. Including ice flow would increase the volume loss, because ice would be transported from higher to lower elevations where the temperatures are warmer. For this reason, we assume that excluding ice flow does not explain why our volume losses are higher than the other studies.~~ Lastly, some of the differences between our study and other published projections could be due models using different ~~the method used to~~ initial~~ise the present-day~~ ice volumes and glaciated areas.

5 Discussion

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. ~~mass balance ,~~One of the notable differences between our study and other global glacier models is that our tuned precipitation lapse rates are very high, for example, 24%/100m for South Asia West and 19%/100m for Central Asia. Other models have used lower precipitation lapse rates (1-2.5%/100m (Huss and Hock 2015), 3%/100m (Marzeion et al. 2012)), but they also bias correct precipitation by multiplying by a scale factor. This scaling factor can be considerably high. Giesen and Oerlemans (2012) found that precipitation needed to be multiplied by a factor of 2.5 to get good agreement with mass balance observations. Radic and Hock (2011) derived, through calibration of present-day mass balance, a precipitation scale factor of as high as 5.6 for Tuyuksu and Golubina glaciers in the Tien Shan. Our lapse rates are high because we do not bias correct the precipitation using a multiplication factor for the present-day. For the future GCM data the gridbox mean precipitation was bias corrected using the ISI-MIP technique. The negative bias that we get when validating the present-day mass balance suggests that snowfall is underestimated in our model. A future study using this model could test whether bias correcting the precipitation before applying the lapse rate correction improves the simulated mass balance.

This is the first attempt to implement a glacier scheme into JULES and so the model has many limitations. One of the key shortcomings 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. These processes could be included in future work by adding a
~~There are several avenues for future model development. Firstly, glacier dynamics in not currently included in the model. Ice flow could be included using a~~ volume-area scaling scheme Bahr- (1997) ~~-~~or a thickness parameterisation based on glacier slope (Marshall and Clarke 2000). Volume-area scaling has been used to model glacier dynamics in coarse resolution (0.5-degree) models where all glaciers in a gridbox are represented by a single ice body (Kotlarski et al. 2010, Hirabayashi, Doll and Kanae 2010). The current configuration of elevated glaciated and unglaciated tiles in JULES makes it well suited to a volume-area scaling model. This would be implemented by growing (shrinking) the elevated glaciated tiles if mass balance is positive (negative) at each elevation band. In the case where the elevated ice tile grows the unglaciated tile would shrink at that elevation band or vice versa.

The volume-area scaling law has been used successfully to model the dynamics of individual glaciers (Radic et al. 2014, Giesen and Oerlemans 2013, Marzeion et al. 2012, Slangen et al. 2014) but has some limitations when applied to coarse models where glaciers are consolidated into a single gridbox. The volume-area scaling law, relates volume to area using a constant scaling exponent which is typically derived from a small sample of glacier observations (Bahr et al. 1997). One of the draw backs is that the law is non-linear, meaning the exponent derived from individual glaciers would overestimate the volume of a large ice grid box such as in our model (Hirabayashi et al. 2013). Furthermore, the exponent may not accurately represent the volume-area relationship in other geographical regions. To overcome these issues a spatially variable scaling exponent could be created using the newly available 0.5-degree data on volume and area contained in the RGI6.

Another shortcoming of the model is the simple treatment of katabatic winds, which is modelled by scaling the synoptic wind speed. ~~A secondly avenue for model development could focus on the treatment of katabatic winds in the model. We represent katabatic winds by simply scaling of the surface wind speed.~~ This could be improved by parameterising katabatic winds based on the gridbox slope and the temperature difference between the glacier surface and the air temperature using the Prandtl model (Oerlemans and Grisogono 2002).

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.

~~There are limitations to the processes that can feasibly be modelled in such a coarse resolution model. For example, avalanching, blowing snow and calving are important processes that affect the local mass balance but would pose a challenge to model on a coarse resolution grid. Finally, w~~While this modelling projects considerable reduction in glacier mass ~~balance~~ for all mountain ranges by the end of this century, it is clear that many of the world's mountain glaciers will evolve in ways that are currently difficult to model. For instance, paraglacial processes during deglaciation lead to enhanced rock falls and debris flows from deglaciating mountain slopes and these deliver rock debris to glacier surfaces. This produces debris-covered glaciers and these are common in many mountain regions, including in Alaska, arid Andes, central Asia and in the Hindu Kush-Himalaya. Thick debris cover (decimetres to metres) limits ice ablation, (e.g., (Lambrecht et al. 2011, Pellicciotti et al.

2014, Lardeux et al. 2016, Rangecroft et al. 2016)) and reverses the mass balance gradient, with comparatively higher ablation rates up glacier than at the debris-covered terminus. This significantly influences glacier dynamics (Benn et al. 2005), and with inefficient sediment evacuation eventually leads to the transition from debris-covered glaciers to rock glaciers (e.g (Monnier and Kinnard 2017). In the context of continued climate warming, the transition from ice glaciers to rock glaciers may enhance the resilience of the mountain cryosphere (Bosson and Lambiel 2016). As a result, better assessment of the response of the mountain cryosphere to climate warming will depend on a clearer understanding of glacier-rock glacier relationships.

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.

6 Conclusions

The first aim of this study was to add a glacier component to JULES to develop a fully integrated model, to simulate the interactions between glacier mass balance, river runoff, water abstraction by irrigation and crop production. To do this we added two new surface types to JULES; elevated glaciated and unglaciated tiles. This allows us to the calculate elevation-dependent mass balance which can be used to study the response of glaciers to climate change. Glacier volume was modelled by growing or shrinking the snowpack, using the elevation-dependent mass balance, but glacier dynamics was not included. Present-day mass balance was calibrated by tuning albedo, wind speed, temperature and precipitation lapse rates to obtain a set of regionally tuned parameters which are then used to model future mass balance. Winter and summer mass balances are reproduced reasonably well for regions where the glaciated area is large, however, the model performs poorly for small glaciers particularly in the summer. The fully integrated model is potentially a useful tool for the scientific community to study the impact of climate change on food and water resources.

The second aim of this study was to make glacier volume projections for the future under a range of high-end climate change scenarios. The ensemble mean volume loss ± one standard deviation is, -64±5% for all glaciers excluding those on the periphery~~peripheral~~ of the Antarctic ice sheet. The small uncertainties in the multi-model mean are caused by the sensitivity of HadGEM3-A to the boundary conditions supplied by the CMIP5 models. Our end of the century global volume loss is 215 ± 20mm, which is higher than values reported by other studies. This is because we used a subset of CMIP5 models with the highest warming levels to drive the model and glacier dynamics is not included which results in more mass loss than other studies that include dynamics. Including parametric uncertainty in the calibration procedure results in an upper bound global volume loss of 281.1 mm sea-level equivalent by the end of the century. The projected ice losses will have an impact on sea-level rise and on water availability in glacier-fed river systems.

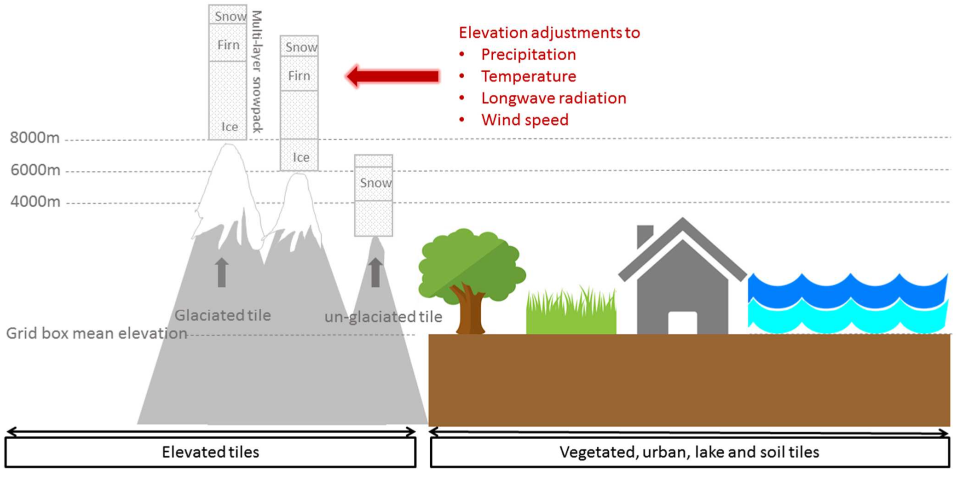

Figure 1 Schematic of JULES surface types inside a single gridbox. The new elevated glaciated and unglaciated tiles are shown on the left-hand side. Note that elevated glaciated and unglaciated tiles are not allowed to share a gridbox with the other tiles.

Table 1 Tuneable parameters for mass balance calculation and their ranges from the literature

Parameter	*Range of values*	*Symbol*	*Units*
Fresh snow albedo (VIS)	0.99 - 0.7	$\alpha_{vis, snow}$	-
Fresh snow albedo (NIR)	0.85 - 0.5	$\alpha_{nir\ snow}$	-
Ice albedo (VIS)	0.7 – 0.1	$\alpha_{vis, ice}$	-
Ice albedo (NIR)	0.6 – 0.1	$\alpha_{nir, ice}$	-
Temperature lapse rate	4 – 9.8	γ_{temp}	^{o}K km^{-1}
Orographic precipitation gradient	5 – 25	γ_{precip}	%/100m
Wind speed scale factor	1 - 4	γ_{wind}	-

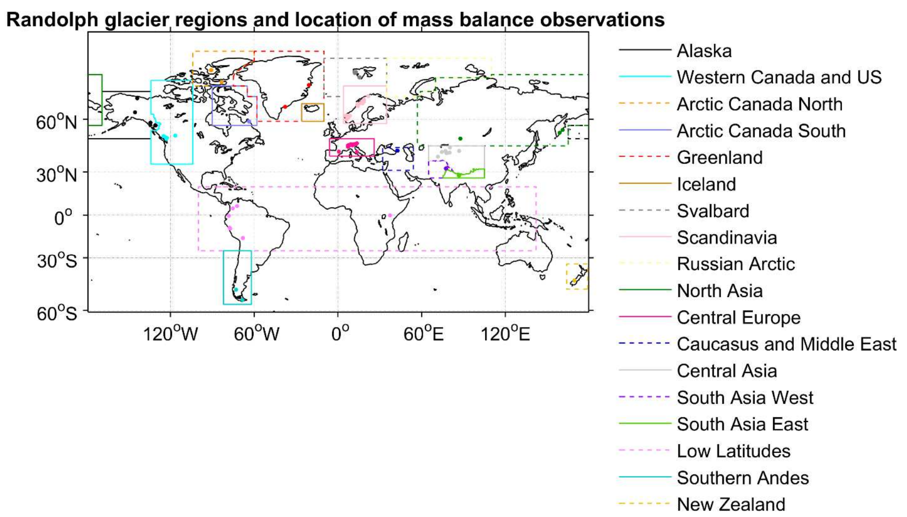

Figure 2: The location of mass balance profile observations glaciers from the World Glacier Monitoring Service and the Randolph Glacier Inventory regions (version 6.0)

Table 2 Best parameter sets for each RGI6 region. The regions are ranked from the lowest to the highest RMSE. There are no observed profiles for Iceland and Russian Artic, so the global mean parameters values are used (bold) for the future simulations

Region	$\alpha_{vis,snow}$	$\alpha_{nir,snow}$	$\alpha_{vis,ice}$	$\alpha_{nir,ice}$	γ_{temp}	γ_{precip}	γ_{wind}
					°K km^{-1}	*%/100m*	
Arctic Canada South	0.94	0.77	0.68	0.53	8.3	16	2.15
Arctic Canada North	0.96	0.70	0.49	0.12	4.2	7	1.10
Greenland	0.95	0.72	0.41	0.19	8.0	15	1.07
Alaska	0.88	0.65	0.56	0.27	8.2	16	1.32
South Asia East	0.91	0.73	0.67	0.56	5.3	9	1.55
South Asia West	0.99	0.73	0.60	0.30	4.0	24	1.69
Western Canada and US	0.97	0.64	0.45	0.26	9.3	8	2.29
Central Asia	0.94	0.74	0.69	0.50	8.1	19	1.40
North Asia	0.94	0.74	0.69	0.50	8.1	19	1.40
Central Europe	0.83	0.63	0.59	0.35	5.8	7	1.83
Svalbard	0.95	0.76	0.54	0.35	9.0	14	1.02
Caucasus and Middle East	0.90	0.71	0.53	0.28	8.3	5	3.32
Scandinavia	0.95	0.76	0.54	0.35	9.0	14	1.02
New Zealand	0.94	0.74	0.69	0.50	8.1	19	1.40
Low Latitudes	0.94	0.74	0.69	0.50	8.1	19	1.40
Southern Andes	0.95	0.76	0.54	0.35	9.0	14	1.02
Mean	**0.93**	**0.72**	**0.58**	**0.37**	**7.55**	**14**	**1.56**

Table 3 Root mean square error (RMSE), correlation coefficient (r), Nash–Sutcliffe efficiency coefficient (NS), mean bias (BIAS) and the number of elevation-band mass balance observations (No Obs) for RGI6 regions. The regions are ranked from the lowest to the highest RMSE.

Region	*RMSE* $m.w.eq.yr^{-1}$	*r*	*NS*	*Bias* $m.w.eq.yr^{-1}$	*No Obs*
Arctic Canada South	0.96	0.61	0.11	0.10	72
Arctic Canada North	1.06	0.19	-0.44	0.52	1,332
Greenland	1.09	0.66	0.14	0.14	90
Alaska	1.36	0.65	0.38	0.06	217
South Asia East	1.41	0.15	-0.34	-0.19	81
South Asia West	1.53	0.62	0.38	-0.09	168
Western Canada and US	1.73	0.69	0.41	-0.40	916
Central Asia	1.81	0.22	-1.15	-0.51	2,519
North Asia	1.95	0.45	-0.04	-0.21	1,335
Central Europe	2.03	0.26	-0.65	0.30	9,561
Svalbard	2.16	0.36	-6.86	-1.21	1,647
Caucasus and Middle East	2.23	0.30	-0.89	0.33	687
Scandinavia	2.40	0.53	0.20	0.67	10,617
New Zealand	2.57	0.58	-0.30	-1.09	45
Low Latitudes	3.06	0.36	-0.71	-0.88	1,016
Southern Andes	3.33	0.26	-12.33	-2.87	118
Global	2.16	0.40	-0.11	0.19	30,421

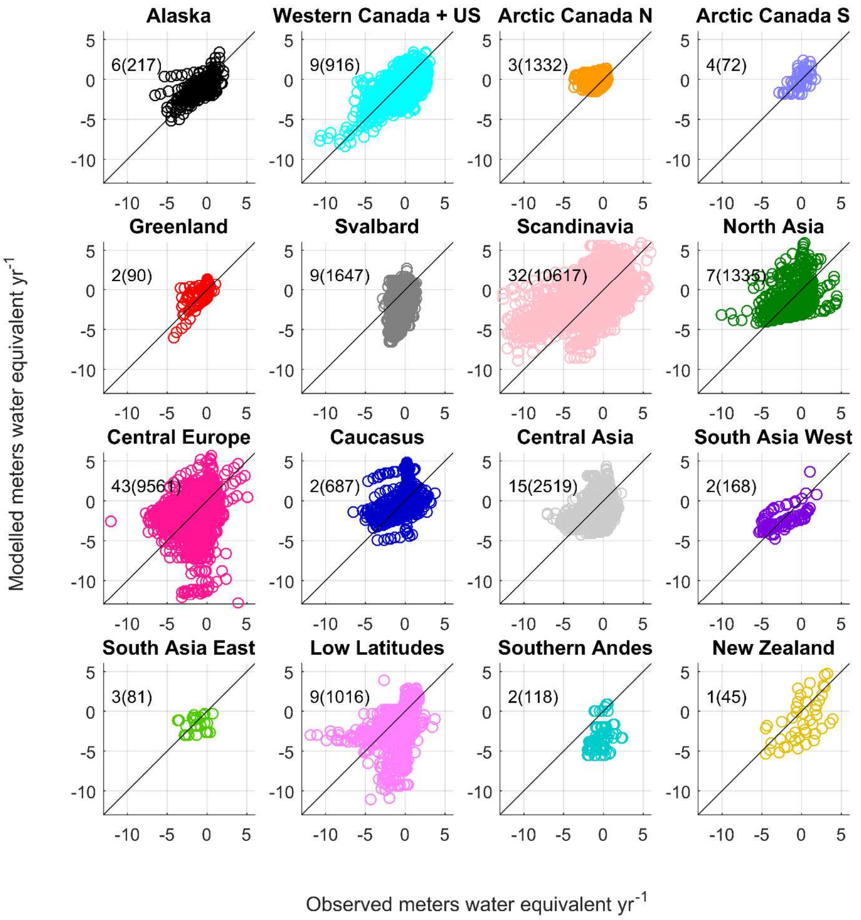

Figure 3: Modelled annual elevation-dependent specific mass balance against observations from the WGMS. The modelled mass balance is simulated on a 0.5-degree grid resolution at 250m elevation bands and the observations are for individual glaciers at elevation levels specific to each glacier. The observed mass balance is interpolated onto the JULES elevation bands. If only a single observation exists, then mass balance for the nearest JULES elevation band is used. The number of glaciers is shown in the top left-hand corner and the number of observation points in brackets.

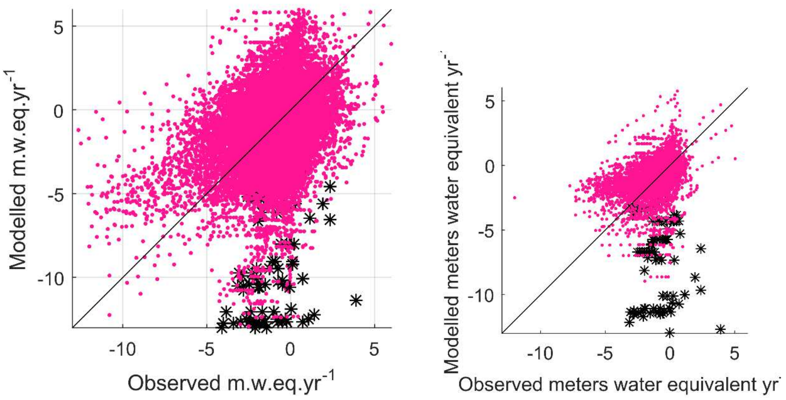

Figure 4: Modelled and observed mass balance profiles for Central Europe (dots) and Maladeta glacier in the Pyrenees (asterixis).

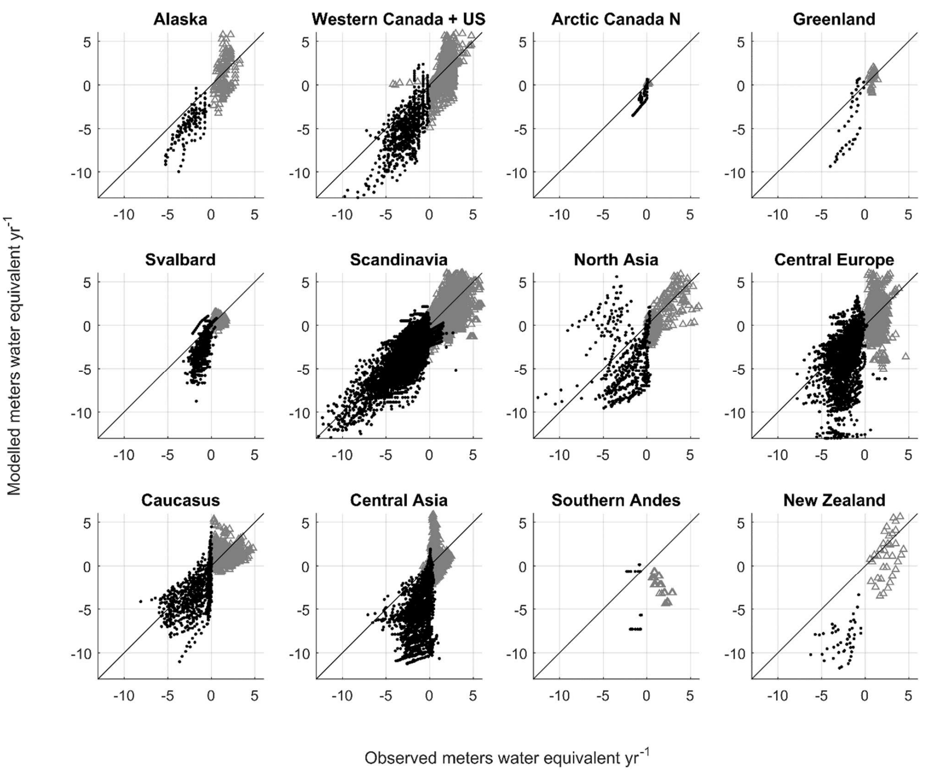

Figure 5 Comparison between modelled and observed elevation-band specific mass balance for winter (grey triangles) and summer (black dots). The modelled mass balance is calculated using the tuned regional parameters from the calibration procedure. ~~In the upper left-hand corner is the summer and winter root mean square error and the number of observations in brackets.~~

Table 4 Winter (bold) and summer (italics) number of elevation-band mass balance observations (No obs), root mean square error (RMSE), correlation coefficient (r), Nash–Sutcliffe efficiency coefficient (NS) and mean bias (BIAS).

Region	*No ~~O~~obs*		*RMSE* *m.w.eq.yr^{-1}*		*r*		*NS*		*BIAS* *m.w.eq.yr^{-1}*	
Alaska	**127**	127	**1.82**	2.43	**0.38**	0.76	**-7.54**	-2.88	**-0.29**	-2.09
Western Canada and US	**767**	729	**1.76**	2.96	**0.53**	0.72	**-2.68**	-2.25	**-0.34**	-2.28
Arctic Canada North	**49**	50	**0.08**	1.09	**0.09**	0.86	**-0.94**	-5.01	**0.04**	-0.79
Greenland	**28**	36	**0.78**	3.45	**0.33**	0.81	**-11.31**	-11.13	**-0.11**	-2.40
Svalbard	**1,122**	1,126	**0.61**	2.25	**0.18**	0.66	**-3.90**	-12.59	**-0.38**	-1.84
Scandinavia	**5,347**	10,679	**1.52**	1.69	**0.61**	0.78	**-0.78**	0.32	**-0.68**	-0.77
North Asia	**854**	828	**1.54**	4.15	**0.71**	0.20	**-0.40**	-3.81	**-1.08**	-2.63
Central Europe	**5,496**	4,804	**1.21**	2.77	**0.12**	0.33	**-5.83**	-4.63	**-0.02**	-1.11
Caucasus + Middle East	**602**	677	**1.39**	2.30	**-0.12**	0.55	**-1.15**	-0.94	**-0.23**	-1.18
Central Asia	**1,778**	1,751	**1.34**	4.87	**0.21**	0.31	**-10.57**	-16.92	**-0.19**	-4.23
Southern Andes	**34**	22	**4.19**	4.11	**-0.81**	-0.08	**-36.73**	-55.59	**-3.81**	-2.36
New Zealand	**45**	45	**3.37**	6.17	**0.42**	0.32	**-10.63**	-17.82	**-0.01**	-5.87
Global	**16,249**	20,874	**1.38**	2.16	**0.49**	0.78	**-1.16**	0.11	**-0.37**	-0.92

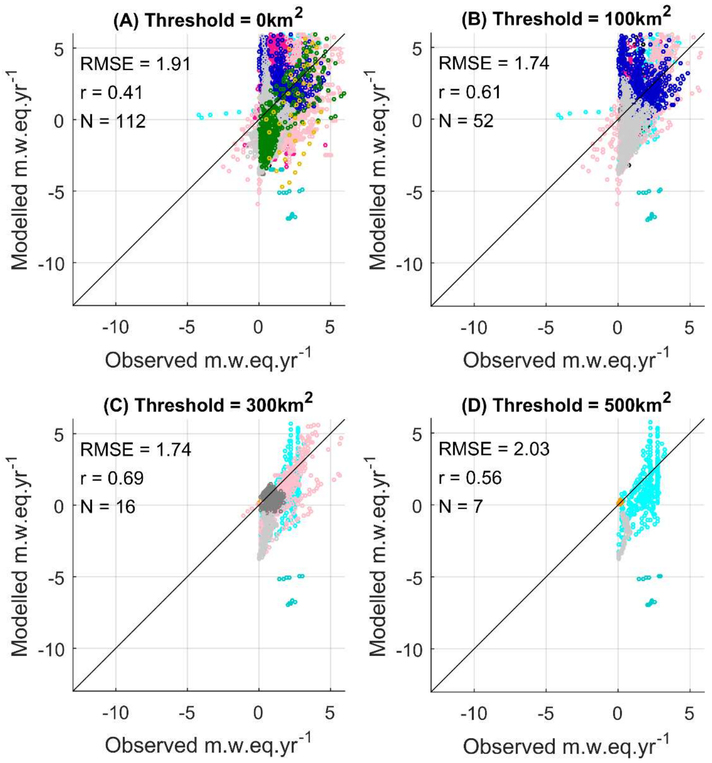

Figure 6 Simulated and observed elevation-dependent winter mass balance when gridboxes with a glacier area of less than 100km^2, 300km^2 and 500km^2 are excluded. The colour identifies the RGI6 regions shown in Figure 2. The RMSE, correlation coefficient and number of glaciers are listed.

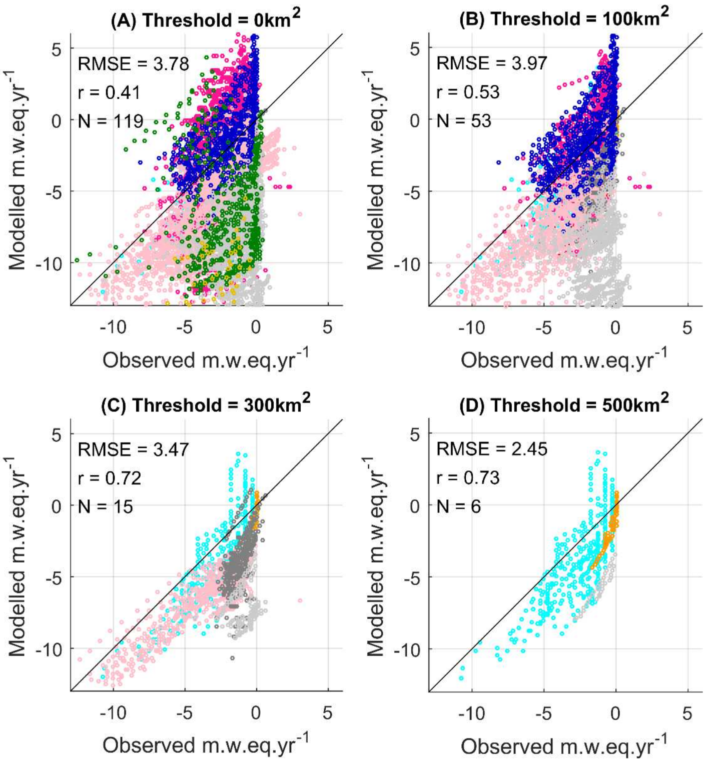

Figure 7 Simulated and observed elevation-dependent summer mass balance when gridboxes with a glacier area of less than 100km^2, 300km^2 and 500km^2 are excluded. The colour identifies the RGI6 regions shown in Figure 2. The RMSE, correlation coefficient and number of glaciers are listed.

Table 5 List of high-end climate change CMIP5 models that are downscaled using HadGEM3-A. The years when the CMIP5 models pass +1.5^{o}C, +2^{o}C and +4^{o}C global average warming relative to the pre-industrial period are shown. *No data is available for 2113 because the bias corrected data ends at 2097.

CMIP5 model	*Ensemble member*	*+ 1.5^{o}C*	*+ 2^{o}C*	*+ 4^{o}C*
IPSL-CM5A-LR	r1i1p1	2015	2030	2068
GFDL-ESM2M	r1i1p1	2040	2055	2113*
HadGEM2-ES	r1i1p1	2027	2039	2074
IPSL-CM5A-MR	r1i1p1	2020	2034	2069
MIROC-ESM-CHEM	r1i1p1	2023	2035	2071
HELIX ACCESS1-0	r1i1p1	2034	2046	2085

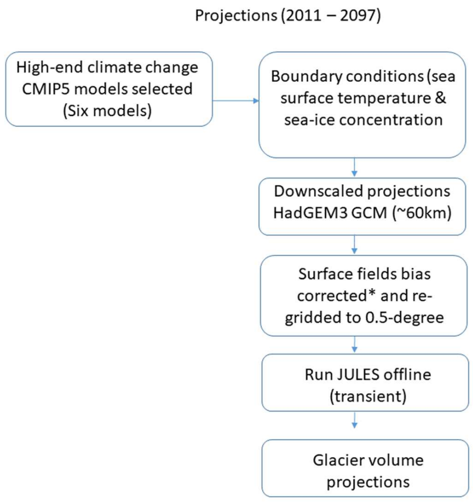

Figure 8 Flow chart showing the experimental set up to calculate future glacier volume. *The bias correction method is described by Hempel et al. (2013b)

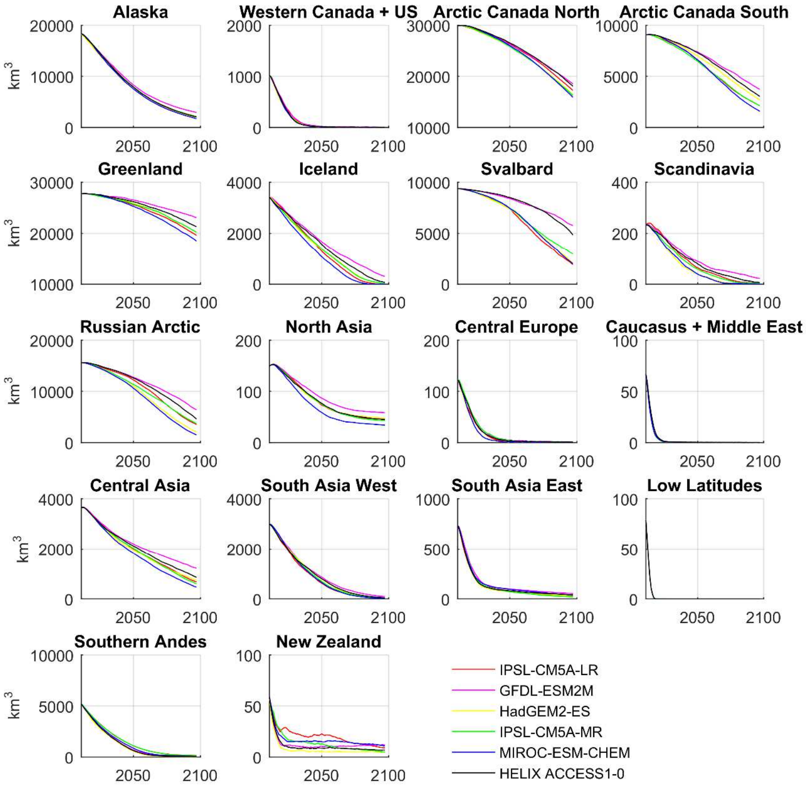

Figure 9 Regional glacier volume projections using the HadGEM3-A ensemble of high-end climate change scenarios.

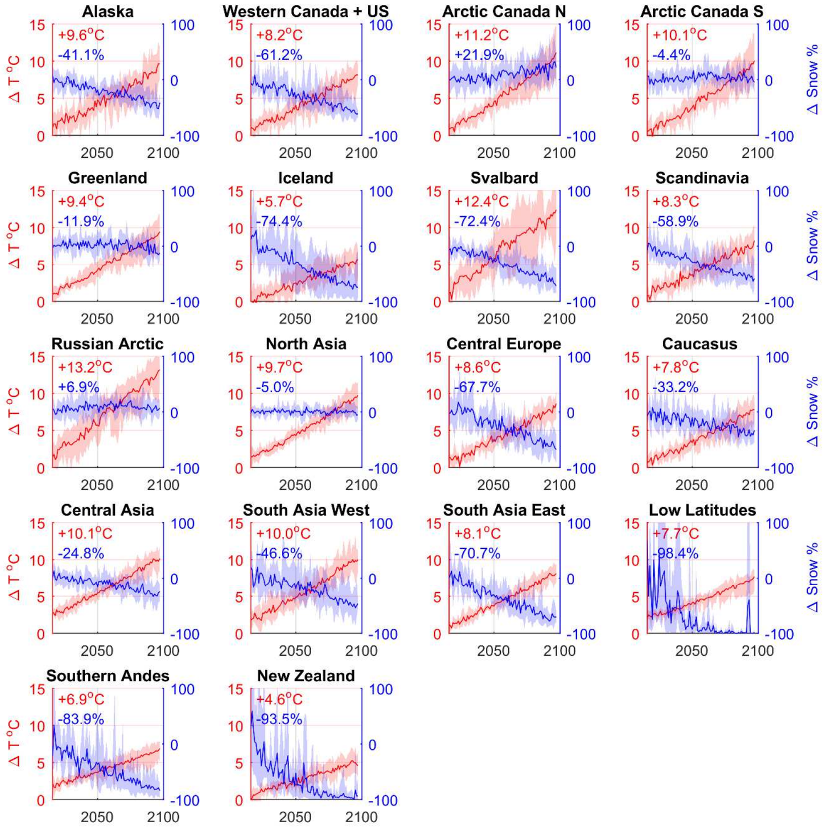

Figure 10 Regional temperature and snowfall changes relative to present day (2011-2015) from the HadGEM3-A ensemble over glaciated grid points. The ensemble mean is shown in the solid line and the range of model projections are shown in the shaded regions.

Table 6 Percentage ice volume loss, relative to the initial volume (ΔV) and ice loss in mm of Sea Level Equivalent (SLE) for the end of the century (2097). Percentage volume loses are shown for low, medium and high elevation ranges as well as for all elevations. The data shows the multi-model mean ± one standard deviation. The conversion of volume to SLE assumes an ocean area of 3.618 x 10^8 km^2. The initial area and volume from the Randolph Glacier Inventory Version 6 is listed in columns 1 and 2.

	Area	Volume	ΔV 0m-9000m	ΔV 0m-2000m	ΔV 2250m-4000m	ΔV 4250m-8000m	SLE
	km^2	km^3	%	%	%	%	mmSLE
Alaska	86,616	19,743	-89±2	-93±1	-55±9	408±18	44.6±1.1
Western Canada and US	14,357	1,070	-100±0	-100±0	-99±0	684±136	2.8±0.0
Arctic Canada North	104,920	32,376	-47±3	-43±4	40±1	-	35.8±3.0
Arctic Canada South	40,861	9,780	-74±8	-72±9	-	-	18.1±2.1
Greenland	126,143	29,856	-31±5	-31±6	37±3	-	20.1±4.4
Iceland	11,052	3,722	-98±3	-98±3	-	-	9.3±0.3
Svalbard	33,932	10,112	-68±16	-65±18	608±158	-	17.0±4.6
Scandinavia	2,948	244	-98±3	-97±3	-92±17	-	0.6±0.0
Russian Arctic	51,552	16,908	-79±10	-77±11	-	-	33.3±4.8
North Asia	2,400	156	-71±5	-97±2	-52±8	220±41	0.3±0.0
Central Europe	2,091	127	-99±0	-100±0	-99±0	-77±24	0.3±0.0
Caucasus & Middle East	1,305	71	-100±0	-100±0	-100±0	-99±0	0.2±0.0
Central Asia	48,415	3,849	-80±7	-	-100±0	-74±9	8.0±0.7
South Asia West	29,561	3,180	-98±1	-	-100±0	-98±1	8.1±0.1
South Asia East	11,148	773	-95±2	-	-100±0	-95±2	1.9±0.0
Low Latitudes	2,341	88	-100±0	-100±0	-100±0	-100±0	0.2±0.0
Southern Andes	29,369	5,701	-98±1	-99±1	-74±14	-57±12	14.4±0.1
New Zealand	1,161	65	-88±5	-100±0	71±62	-	0.1±0.0
Global	600,172	137,821	-64±5	-61±6	-36±3	-84±5	215.2±21.3

	~~Area~~	~~Volume~~	~~ΔV~~	~~SLE~~
	~~km^2~~	~~km^3~~	~~2011-2097~~	~~2011-2097~~
~~Alaska~~	~~86,616~~	~~19,743~~	~~-89±2~~	~~44.6±1.1~~
~~Western Canada and US~~	~~14,357~~	~~1,070~~	~~-100±0~~	~~2.8±0.0~~
~~Arctic Canada North~~	~~104,920~~	~~32,376~~	~~-47±3~~	~~35.8±3.0~~
~~Arctic Canada South~~	~~40,861~~	~~9,780~~	~~-74±8~~	~~18.1±2.1~~
~~Greenland~~	~~126,143~~	~~29,856~~	~~-31±5~~	~~20.1±4.4~~
~~Iceland~~	~~11,052~~	~~3,722~~	~~-98±3~~	~~9.3±0.3~~
~~Svalbard~~	~~33,932~~	~~10,112~~	~~-68±16~~	~~17.0±4.6~~
~~Scandinavia~~	~~2,948~~	~~244~~	~~-98±3~~	~~0.6±0.0~~
~~Russian Arctic~~	~~51,552~~	~~16,908~~	~~-79±10~~	~~33.3±4.8~~

~~*North Asia*~~	~~2,400~~	~~156~~	~~-71±5~~	~~0.3±0.0~~
~~*Central Europe*~~	~~2,091~~	~~127~~	~~-99±0~~	~~0.3±0.0~~
~~*Caucasus and Middle East*~~	~~1,305~~	~~71~~	~~-100±0~~	~~0.2±0.0~~
~~*Central Asia*~~	~~48,415~~	~~3,849~~	~~-80±7~~	~~8.0±0.7~~
~~*South Asia West*~~	~~29,561~~	~~3,180~~	~~-98±1~~	~~8.1±0.1~~
~~*South Asia East*~~	~~11,148~~	~~773~~	~~-95±2~~	~~1.9±0.0~~
~~*Low Latitudes*~~	~~2,341~~	~~88~~	~~-100±0~~	~~0.2±0.0~~
~~*Southern Andes*~~	~~29,369~~	~~5,701~~	~~-98±1~~	~~14.4±0.1~~
~~*New Zealand*~~	~~1,161~~	~~65~~	~~-88±5~~	~~0.1±0.0~~
~~*Global*~~	~~600,172~~	~~137,821~~	~~-64±5~~	~~215.2±21.3~~

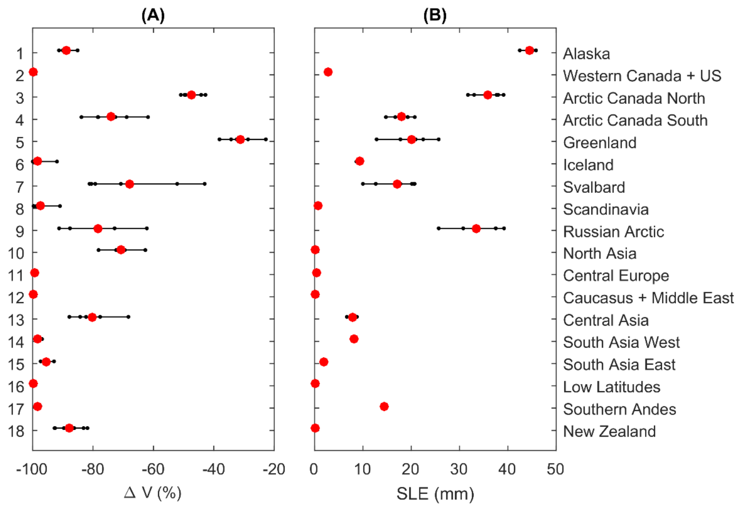

Figure 11 (A) Regional percentage volume losses at the end of the century (2097), relative to the initial volume and (B) volume losses expressed in sea-level equivalent contributions. The large red dots represent the multi-model mean and the small black dots are the individual HadGEM3-A model runs.

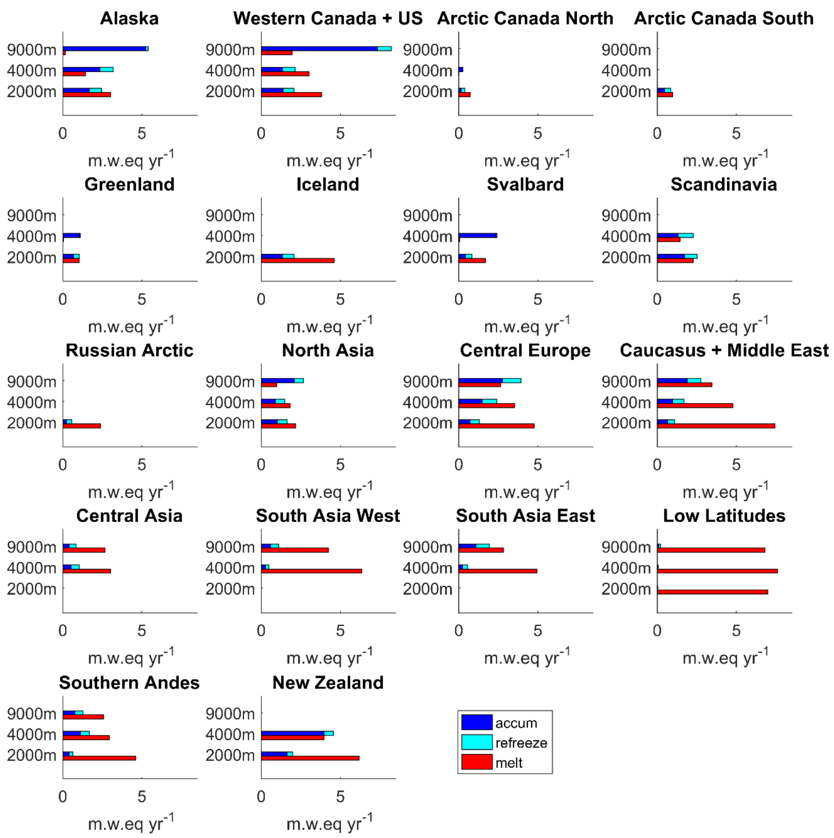

Figure 12 Modelled annual surface mass balance components; accumulation, refreezing and melting for the period 1980-2000 for RGI6 regions. To make the figure easier to read melting is given as positive sign and sublimation is excluded because its contribution is very small. Mass balance components are averaged over low (0-2000m), medium (2250m-4000m) and high (4250-9000m) elevation ranges.

Formatted: Keep with next

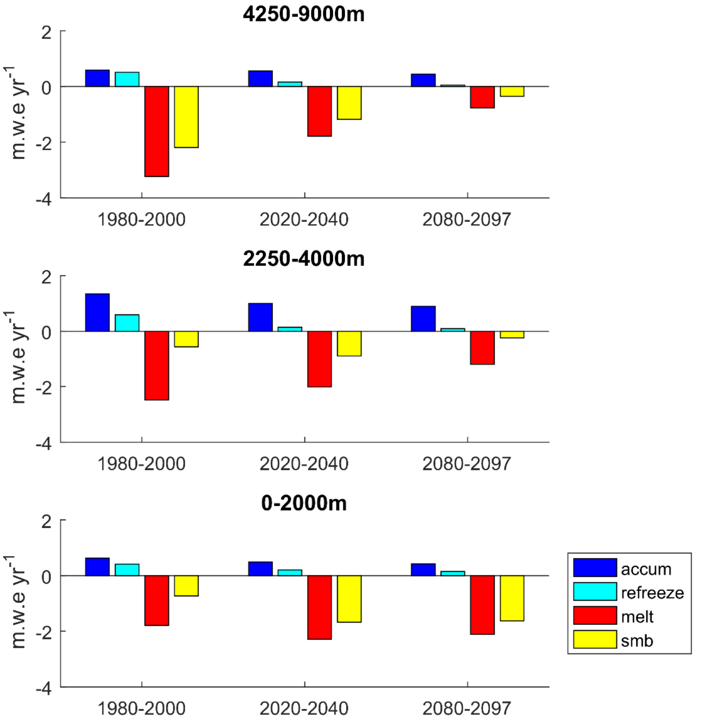

Figure 13 Global mass balance components for three elevation ranges. The historical period is calculated using the WDFEI data and the future period is the multi-model means of all GCMs. The bars show the averages over the time periods for accumulation, refreezing, melt and mass balance rates.

Formatted: Keep with next

Formatted: Caption

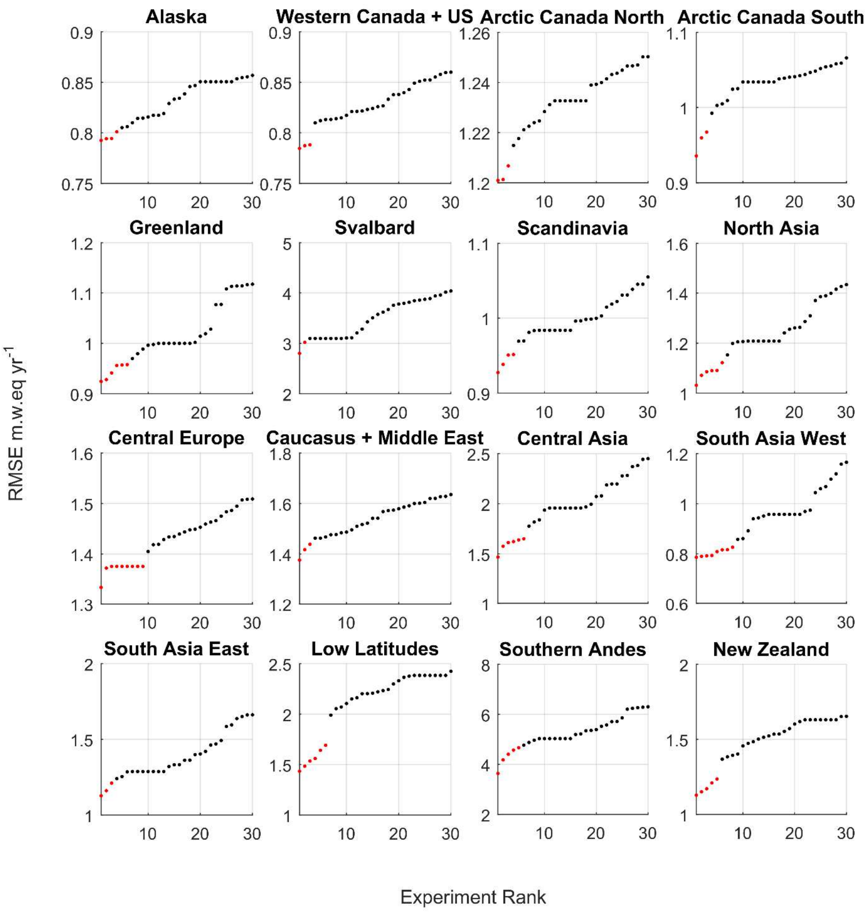

Figure 14 Calibration experiments raked according the root mean square error between simulated and observed mass balance profiles for RGI6 regions. There are 198 experiments but only the top 30 have been plotted to make the figure easier to read. The red dots indicate experiments that perform equally well.

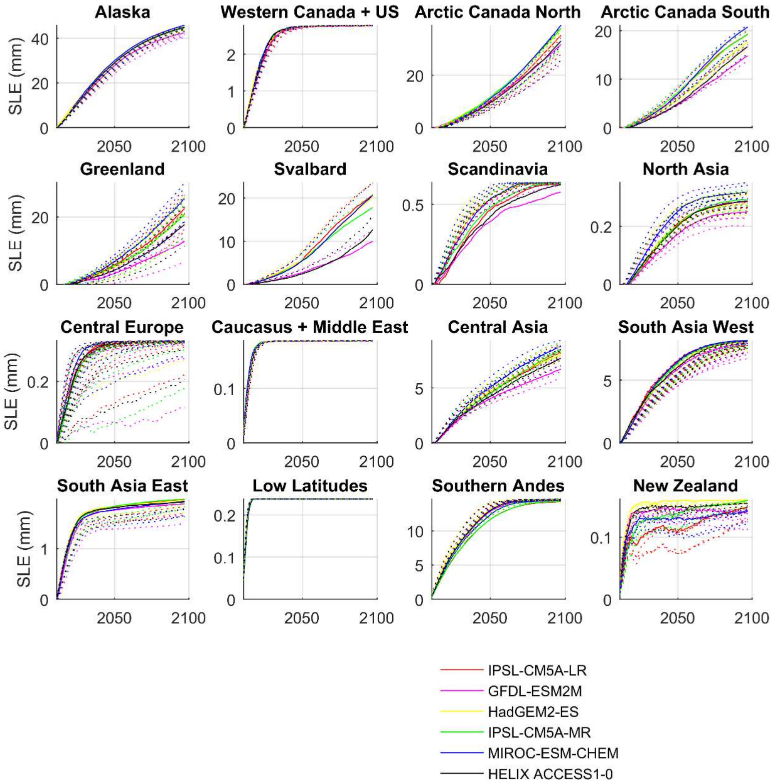

Figure 15 Regional volume losses expressed in sea-level equivalent including parametric uncertainty in mass balance parameters. The solid lines show the volume loss for each downscaled CMIP5 GCM using the optimum parameter sets. The dashed lines are for runs which use other equally 'good' parameter sets based on the RMSE.

Table 7 Regional ensemble mean, minimum and maximum volume losses for 2097 in sea-level equivalent (mm) when the present-day mass balance is calibrated in different ways. Columns 1-3; mass balance is calibrated by minimising the RMSE. Columns 4-6; mass balance is calibrated using an ensemble of equally plausible RMSE values. Columns 7-9; mass balance is calibrated by minimising the RMSE, minimising the bias and maximising the correlation coefficient. ~~when mass balance is cabraring by minimising RMSE, using equally plausible RMSE and using RMSE, Volume losses when calibration parameters are selebed ny minimising the RMSE Min RMSE, (The first three columns using the optimum parameter set for mass balance and the last three columns are for runs which include parametric uncertainty.~~

	Optimum parameter			Equally plausible RMSE			Extra performance metrics		
	SLE_{mean}	SLE_{min}	SLE_{max}	SLE_{mean}	SLE_{min}	SLE_{max}	SLE_{mean}	SLE_{min}	SLE_{max}
Alaska	44.6	42.5	45.8	43.8	40.5	45.8	43.6	38.2	46.3
Western Canada and US	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8
Arctic Canada North	35.8	31.8	39.1	32.8	24.3	39.1	37.2	22.3	61.8
Arctic Canada South	18.1	14.8	20.8	17.9	13.7	21.1	20.3	14.8	24.1
Greenland	20.1	12.9	25.7	20.4	6.7	30.2	23.5	14.0	31.8
Iceland	9.3	8.7	9.5	9.3	8.7	9.5	9.4	8.5	9.5
Svalbard	17.0	10.0	20.7	18.4	10.0	23.6	19.7	10.0	25.8
Scandinavia	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
Russian Arctic	33.3	25.7	39.2	33.3	25.7	39.2	36.6	25.1	42.8
North Asia	0.3	0.3	0.3	0.3	0.2	0.4	0.3	0.3	0.4
Central Europe	0.3	0.3	0.3	0.3	0.1	0.3	0.3	0.2	0.3
Caucasus and Middle East	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Central Asia	8.0	6.7	8.8	8.0	5.9	9.3	8.1	5.9	9.5
South Asia West	8.1	8.0	8.2	7.8	6.7	8.2	7.9	7.1	8.2
South Asia East	1.9	1.9	2.0	1.8	1.5	2.0	1.8	1.6	2.0
Low Latitudes	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Southern Andes	14.4	14.2	14.5	14.4	14.2	14.6	14.4	14.2	14.6
New Zealand	0.1	0.1	0.2	0.1	0.1	0.2	0.1	0.1	0.2
Global	215.2	181.5	238.9	212.6	162.2	247.3	227.1	166.1	281.1
Global SLE_{max} - SLE_{min}	57.3			85.1			115.0		

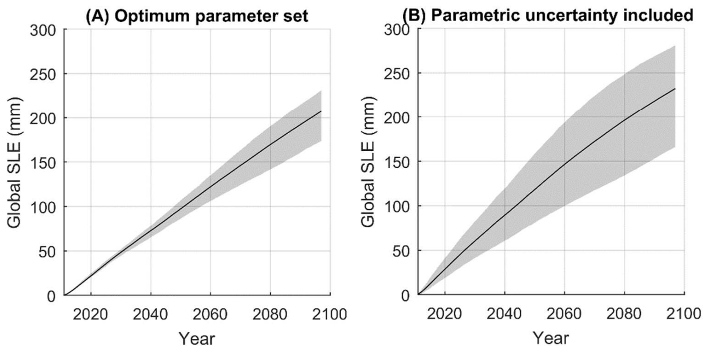

Figure 16 Multi-model mean (black line) and ensemble spread (shaded) global volume loss in sea-level equivalent. (A) is the volume loss when optimum parameters sets are selected by minimising the RMSE and (B) is volume loss when optimum parameters sets are selected using additional performance metrics (minimising RMSE, minimising the bias and maximising the correlation coefficient).

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

Formatted: Normal

Acknowledgements

The research leading to these results was funded by from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement n° 603864. We would like to thank the Natural Environment Research Council (NERC) for the use of the -Joint Analysis System Meeting Infrastructure Needs (JASMIN) super computer cluster.

Formatted: Font: Times New Roman, 10 pt, Font color: Auto, Pattern: Clear

References

~~*WGMS 2017, Global Glacier Change Bulletin No. 2 (2014-2015).* Zurich, Switzerland.~~

Ageta, Y. & K. Higuchi (1984) ESTIMATION OF MASS BALANCE COMPONENTS OF A SUMMER-ACCUMULATION TYPE GLACIER IN THE NEPAL HIMALAYA. *Geografiska Annaler Series a-Physical Geography,* 66**,** 249-255.

Bahr, D. B., M. F. Meier & S. D. Peckham (1997) The physical basis of glacier volume-area scaling. *Journal of Geophysical Research-Solid Earth,* 102**,** 20355-20362.

Bell, V. A., A. L. Kay, R. G. Jones & R. J. Moore (2007) Development of a high resolution grid-based river flow model for use with regional climate model output. *Hydrology and Earth System Sciences,* 11**,** 532-549.

Benn, D. I., M. Kirkbride, L. A. Owen & V. Brazier. 2005. *Glaciated Valley Landsystems, Hodder Education*. LondonLondon: D.J.A. Evans (Ed.).

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.

Bosson, J. B. & C. Lambiel (2016) Internal Structure and Current Evolution of Very Small Debris-Covered Glacier Systems Located in Alpine Permafrost Environments. *Frontiers in Earth Science,* 4.

Bruland, O. & J. O. Hagen (2002) Glacial mass balance of Austre Broggerbreen (Spitsbergen), 1971-1999, modelled with a precipitation-run-off model. *Polar Research,* 21**,** 109-121.

Chen, X. L. & T. J. Zhou (2015) Distinct effects of global mean warming and regional sea surface warming pattern on projected uncertainty in the South Asian summer monsoon. *Geophysical Research Letters,* 42**,** 9433-9439.

Farinotti, D., M. Huss, A. Bauder, M. Funk & M. Truffer (2009) A method to estimate the ice volume and ice-thickness distribution of alpine glaciers. *Journal of Glaciology,* 55**,** 422-430.

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.

Giesen, R. H. & J. Oerlemans (2013) Climate-model induced differences in the 21st century global and regional glacier contributions to sea-level rise. *Climate Dynamics,* 41**,** 3283-3300.

Gohar, L. K., J. A. Lowe & D. Bernie (2017) The Impact of Bias Correction and Model Selection on Passing Temperature Thresholds. *Journal of Geophysical Research-Atmospheres,* 122**,** 12045-12061.

Grenfell, T. C., S. G. Warren & P. C. Mullen (1994) REFLECTION OF SOLAR-RADIATION BY THE ANTARCTIC SNOW SURFACE AT ULTRAVIOLET, VISIBLE, AND NEAR-INFRARED WAVELENGTHS. *Journal of Geophysical Research-Atmospheres,* 99**,** 18669-18684.

Greuell, W. & T. Konzelmann (1994) NUMERICAL MODELING OF THE ENERGY-BALANCE AND THE ENGLACIAL TEMPERATURE OF THE GREENLAND ICE-SHEET - CALCULATIONS FOR THE ETH-CAMP LOCATION (WEST GREENLAND, 1155M ASL). *Global and Planetary Change,* 9**,** 91-114.

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.

Hadley, O. L. & T. W. Kirchstetter (2012) Black-carbon reduction of snow albedo. *Nature Climate Change,* 2**,** 437-440.

Hempel, S., K. Frieler, L. Warszawski, J. Schewe & F. Piontek (2013) A trend-preserving bias correction – the ISI-MIP approach. *Earth System Dynamics,* 4**,** 219-236.

Hirabayashi, Y., P. Doll & S. Kanae (2010) Global-scale modeling of glacier mass balances for water resources assessments: Glacier mass changes between 1948 and 2006. *Journal of Hydrology,* 390**,** 245-256.

Hirabayashi, Y., Y. Zang, S. Watanabe, S. Koirala & S. Kanae (2013) Projection of glacier mass changes under a high-emission climate scenario using the global glacier model HYOGA2. *Hydrological Research Letters,* 7**,** 6-11.

Hock, R. (2005) Glacier melt: a review of processes and their modelling. *Progress in Physical Geography,* 29**,** 362-391.

Hudson, S. R., S. G. Warren, R. E. Brandt, T. C. Grenfell & D. Six (2006) Spectral bidirectional reflectance of Antarctic snow: Measurements and parameterization. *Journal of Geophysical Research-Atmospheres,* 111.

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

Huss, M., M. Funk & A. Ohmura (2009) Strong Alpine glacier melt in the 1940s due to enhanced solar radiation. *Geophysical Research Letters,* 36**,** L23501.

Huss, M. & R. Hock (2015) A new model for global glacier change and sea-level rise. *Frontiers in Earth Science,* 3**,** 54.

Huss, M. & R. Hock (2018) Global-scale hydrological response to future glacier mass loss. *Nature Climate Change,* 8**,** 135-140.

Immerzeel, W. W. P., F. Bierkens, M. F. P. (2013) Rising river flows throughout the twenty-first century in two Himalayan glacierized watersheds. *Nature Geoscience,* 3**,** 742–745

Kaser, G., J. G. Cogley, M. B. Dyurgerov, M. F. Meier & A. Ohmura (2006) Mass balance of glaciers and ice caps: Consensus estimates for 1961-2004. *Geophysical Research Letters,* 33.
Kotlarski, S., D. Jacob, R. Podzun & F. Paul (2010) Representing glaciers in a regional climate model. *Climate Dynamics,* 34**,** 27-46.
Lambrecht, A., C. Mayer, W. Hagg, V. Popovnin, A. Rezepkin, N. Lomidze & D. Svanadze (2011) A comparison of glacier melt on debris-covered glaciers in the northern and southern Caucasus. *Cryosphere,* 5**,** 525-538.
Landolt-Bornstein (1987) Numerical Data and Functional Relationships in Science and Technology. *Group V Geophysics and Space Research, Physical and Chemical Properties of Air, p 35,* 4B Meteorology.
Lardeux, P., N. Glasser, T. Holt & B. Hubbard (2016) Glaciological and geomorphological map of Glacier Noir and Glacier Blanc, French Alps. *Journal of Maps,* 12**,** 582-596.
Lavielle, M. (2005) Using penalized contrasts for the change-point problem. *Signal Processing,* 85**,** 1501–1510.
Lutz, A. F., W. W. Immerzeel, A. B. Shrestha & M. F. P. Bierkens (2014) Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation. *Nature Climate Change,* 4**,** 587–592.
Marshall, S. J. & G. K. C. Clarke (2000) Ice sheet inception: subgrid hypsometric parameterization of mass balance in an ice sheet model (vol 15, pg 533, 1999). *Climate Dynamics,* 16**,** 319-319.
Marzeion, B., A. H. Jarosch & J. M. Gregory (2014) Feedbacks and mechanisms affecting the global sensitivity of glaciers to climate change. *Cryosphere,* 8**,** 59-71.
Marzeion, B., A. H. Jarosch & M. Hofer (2012) Past and future sea-level change from the surface mass balance of glaciers. *The Cryosphere Discuss.,* 6**,** 3177-3241.
McKay, M. D., R. J. Beckman & W. J. Conover (1979) Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics,* 21**,** 239-245.
Meier, M. F., M. B. Dyurgerov, U. K. Rick, S. O'Neel, W. T. Pfeffer, R. S. Anderson, S. P. Anderson & A. F. Glazovsky (2007) Glaciers Dominate Eustatic Sea-Level Rise in the 21st Century. *Science,* 317**,** 1064-1067.
Monnier, S. & C. Kinnard (2017) Pluri-decadal (1955-2014) evolution of glacier-rock glacier transitional landforms in the central Andes of Chile (30-33 degrees S). *Earth Surface Dynamics,* 5**,** 493-509.
Oerlemans, J. & B. Grisogono (2002) Glacier winds and parameterisation of the related surface heat fluxes. *Tellus Series a-Dynamic Meteorology and Oceanography,* 54**,** 440-452.
Oki, T., T. Nishimura, and P. A. Dirmeyer (1999) Assessment of annual runoff from land surface models using Total Runoff Integrating Pathways (TRIP). *Journal of the Meteorological Society of Japan,* 77**,** 135–255.
Osborne, T. G., J. Hooker, J. Williams, K. Wiltshire, A. Betts, R. Wheeler, T. (2014) JULES-crop: a parametrisation of crops in the Joint UK Land Environment Simulator. *Geosci. Model Dev. Discuss.,* 7.
Pellicciotti, F., M. Carenzo, R. Bordoy & M. Stoffel (2014) Changes in glaciers in the Swiss Alps and impact on basin hydrology: Current state of the art and future research. *Science of the Total Environment,* 493**,** 1152-1170.
Radic, V., A. Bliss, A. C. Beedlow, R. Hock, E. Miles & J. G. Cogley (2014) Regional and global projections of twenty-first century glacier mass changes in response to climate scenarios from global climate models. *Climate Dynamics,* 42**,** 37-58.
Radic, V. & R. Hock (2011) Regionally differentiated contribution of mountain glaciers and ice caps to future sea-level rise. *Nature Geosci,* 4**,** 91-94.
Raftery, A. E., A. Zimmer, D. M. W. Frierson, R. Startz & P. R. Liu (2017) Less than 2 degrees C warming by 2100 unlikely. *Nature Climate Change,* 7**,** 637.
Rangecroft, S., A. J. Suggitt, K. Anderson & S. Harrison (2016) Future climate warming and changes to mountain permafrost in the Bolivian Andes. *Climatic Change,* 137**,** 231-243.
Rasmussen, R., B. Baker, J. Kochendorfer, T. Meyers, S. Landolt, A. P. Fischer, J. Black, J. M. Theriault, P. Kucera, D. Gochis, C. Smith, R. Nitu, M. Hall, K. Ikeda & E. Gutmann (2012) HOW WELL ARE WE MEASURING SNOW? The NOAA/FAA/NCAR Winter Precipitation Test Bed. *Bulletin of the American Meteorological Society,* 93**,** 811-829.
Rebecca, K., P. Fearnhead & I. A. Eckley (2012) Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association,* 107**,** 1590–1598.
Reijmer, C. H., R. Bintanja & W. Greuell (2001) Surface albedo measurements over snow and blue ice in thematic mapper bands 2 and 4 in Dronning Maud Land, Antarctica. *Journal of Geophysical Research-Atmospheres,* 106**,** 9661-9672.
RGI Consortium. 2017. Randolph Glacier Inventory – A Dataset of Global
Glacier Outlines: Version 6.0. In *Global Land Ice Measurements from Space*. Colorado, USA, Digital Media.
Roesch, A., M. Wild, R. Pinker & A. Ohmura (2002) Comparison of spectral surface albedos and their impact on the general circulation model simulated surface climate. *Journal of Geophysical Research-Atmospheres,* 107.
Rogelj, J., M. den Elzen, N. Hohne, T. Fransen, H. Fekete, H. Winkler, R. S. Chaeffer, F. Ha, K. Riahi & M. Meinshausen (2016) Paris Agreement climate proposals need a boost to keep warming well below 2 degrees C. *Nature,* 534**,** 631-639.
Singh, P. 2001. *Snow and Glacier Hydrology*. Springer Netherlands.
Slangen, A. B. A., M. Carson, C. A. Katsman, R. S. W. van de Wal, A. Kohl, L. L. A. Vermeersen & D. Stammer (2014) Projecting twenty-first century regional sea-level changes. *Climatic Change,* 124**,** 317-332.

Sorg, A., T. Bolch, M. Stoffel, O. Solomina & M. Beniston (2012) Climate change impacts on glaciers and runoff in Tien Shan (Central Asia). *Nature Climate Change,* 2**,** 725-731.

Sorg, A., M. Huss, M. Rohrer & M. Stoffel (2014) The days of plenty might soon be over in glacierized Central Asian catchments. *Environmental Research Letters,* 9.

Walters, D., I. Boutle, M. Brooks, T. Melvin, R. Stratton, S. Vosper, H. Wells, K. Williams, NigelWood, T. Allen, A. Bushell, D. Copsey, P. Earnshaw, J. Edwards, M. Gross, S. Hardiman, C. Harris, J. Heming, N. Klingaman, R. Levine, J. Manners, G. Martin, S. Milton, M. Mittermaier, C. Morcrette, T. Riddick, M. Roberts, C. Sanchez, P. Selwood, A. Stirling, C. Smith, D. Suri, W. Tennant, P. L. Vidale, JonathanWilkinson, M. Willett, S. Woolnough & P. Xavier (2017) The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. *Geoscientific Model Development,* 10**,** 1487-1520.

Warren, S. G. & W. J. Wiscombe (1980) A model for the spectral albedo of snow, 1. Pure snow. *Journal of Atmospheric Science,* 37**,** 2712-2733.

Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. & P. Viterbo (2014) The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*.

Weedon, G. P., S. Gomes, P. Viterbo, W. J. Shuttleworth, E. Blyth, H. Osterle, J. C. Adam, N. Bellouin, O. Boucher & M. Best (2011) Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century. *Journal of Hydrometeorology,* 12**,** 823-848.

WGMS 2017, Global Glacier Change Bulletin No. 2 (2014-2015). Zurich, Switzerland.

Zhang, Y. L., S. C. Kang, Z. Y. Cong, J. Schmale, M. Sprenger, C. L. Li, W. Yang, T. G. Gao, M. Sillanpaa, X. F. Li, Y. J. Liu, P. F. Chen & X. L. Zhang (2017) Light-absorbing impurities enhance glacier albedo reduction in the southeastern Tibetan plateau. *Journal of Geophysical Research-Atmospheres,* 122**,** 6915-6933.

Formatted: Indent: Left: 0 cm, First line: 0 cm

Page 53: [1] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [2] Formatted Table	Sarah Shannon	02/08/2018 17:26:00

Formatted Table

Page 53: [3] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [4] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [5] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [6] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [7] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [8] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [9] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [10] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [11] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [12] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt, Superscript

Page 53: [12] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt, Superscript

Page 53: [13] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt, Superscript

Page 53: [13] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt, Superscript

Page 53: [14] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [15] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [16] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [17] Formatted Sarah Shannon 02/08/2018 17:26:00

Centered

Page 53: [18] Formatted Sarah Shannon 02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [19] Formatted Sarah Shannon 02/08/2018 17:26:00

Centered

Page 53: [20] Formatted Sarah Shannon 02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [21] Formatted Sarah Shannon 02/08/2018 17:26:00

Centered

Page 53: [22] Formatted Sarah Shannon 02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [23] Formatted Sarah Shannon 02/08/2018 17:26:00

Centered

Page 53: [24] Formatted Sarah Shannon 02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [25] Formatted Sarah Shannon 02/08/2018 17:26:00

Centered

Page 53: [26] Formatted Sarah Shannon 02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [27] Formatted Sarah Shannon 02/08/2018 17:26:00

Centered

Page 53: [28] Formatted Sarah Shannon 02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [29] Formatted Sarah Shannon 02/08/2018 17:26:00

Centered

Page 53: [30] Formatted Sarah Shannon 02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [31] Formatted Sarah Shannon 02/08/2018 17:26:00

Centered

Page 53: [32] Formatted Sarah Shannon 02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [33] Formatted Sarah Shannon 02/08/2018 17:26:00

Centered

Page 53: [34] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [35] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [36] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [36] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [37] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [38] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [39] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [40] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [41] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [42] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [43] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [44] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [45] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [46] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [47] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [48] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [49] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [50] Formatted	Sarah Shannon	02/08/2018 17:22:00

Font: (Default) +Headings (Times New Roman), 10 pt

Page 53: [51] Formatted	Sarah Shannon	02/08/2018 17:26:00

Centered

Page 53: [52] Formatted Table	Sarah Shannon	02/08/2018 17:26:00

Formatted Table