egusphere-2024-863 response to Anonymous Referee #1

We thank anonymous referee #1 (R1) for the encouraging and thorough review of our manuscript with clear recommendations. We have considered our response to each comment carefully, especially to those that require clarification or are critical for which we provide more detailed responses with clear justifications. We did find that a small number of comments stem from some confusion about the underlying methodology of this study (e.g. those related the WRF albedo). We take full responsibility for this and plan to be more explicit about some of the details in the revised manuscript. Furthermore, these were still very valuable comments to us and underpin a number of improvements that we propose below. We have included all of the original referee comments in black. Our responses are in blue. Any proposed revisions are written with **bold underlined text**. We look forward to the response from the editor at their earliest convenience.

In this paper, the authors combine the JULES model, which solves the glacier energy balance, with OGGM, which simulates ice flow, and investigate the mass balance over 500 tropical glaciers in the Vilcanota-Urubamba basin in Peru. Specifically, the authors calibrate the surface energy balance model parameters based on 30 glaciers during the period of 2000-2018 and project the mass balance for all glaciers in the region until 2100 with RCP4.5 and RCP8.5. They find that sublimation plays a minor role in glacier evolution at the basal scale, and their mass balance projections are more conservative than previous models. For example, the JULES-OGGM model estimates that 17% of the ice mass will remain by 2100 under RCP4.5, while other models (GloGEM and MAR2012) predict this number to be only 2%.

The manuscript is well written and provides a valuable addition to the current literature of physics-based glacier models, and glacier modeling in general in the region of the tropical Andes. However, I believe that the authors need to provide more information on the climatic input data used in the surface energy balance model. I also suggest that the authors add a sensitivity study for the input data (especially for albedo). Comments and suggestions are given in the list below.

Major comments:

JULES model input parameters: One of the main challenges when using surface energy balance modeling is obtaining the required input data. The authors mention that the data for the other meteorological variables (other than temperature and precipitation) were “generated by resampling (repeating) the 1980-2018 WRF simulations to produce a continuous 2019-2100 time series” (lines 131-133). By not using predictions from CMIP5 for these variables (radiation, relative humidity, wind speed, etc.), the authors are not using future climate information, but rather (randomly?) resampling the current climate for those variables. Hence, I am wondering about large biases in their model input data for the surface balance model. Can the authors please elaborate on their reasoning and discuss biases?

We completely agree that establishing suitable climate input data is an important and challenging aspect of any surface energy balance modelling study and we put considerable effort into discussing and developing a justifiable approach to do this. As with the temperature and precipitation variables, it was clear that we could not simply apply the raw CMIP5 climate simulations given the inherent biases in the climate model outputs, in part due to the coarser grid spacing of CMIP5 climate models compared to our WRF model. While statistical downscaling methods, such as the quantile delta mapping approach used in this study, have been applied widely to precipitation and temperature variables, they are not routinely used for other variables such as wind and radiation. This is due to the weaker direct link between these variables across spatial scales. For example, a future increase in meridional near-surface wind in the CMIP5 models does not necessarily translate into a future increase in near-surface wind at the WRF model resolution, due to the influence of topography (amongst others). Similarly, radiation change may depend more on local temperature and humidity than large-scale radiation change. For precipitation and temperature, we had the additional benefit of having a wide distribution of meteorological stations across the catchment to validate our downscaling approaches (more information provided in Potter et al., 2023). The same cannot be said for other meteorological variables which are not routinely monitored in the catchment. On balance, we felt that attempting to downscale the CMIP5 projections of the other variables could introduce additional biases in our input data than resampling our historical, dynamically-downscaled, data. We appreciate, however, that the opposite argument could be made i.e. that by using our historical data, we could be missing important changes in these other variables. Of these, our study indicates that the radiation balance is most important. To explore this aspect in more detail, we extracted the incident long and shortwave radiation data from each CMIP5 model at the point closest to the middle of the domain (at -13.69,-72.02). We found that annual average incident longwave radiation is projected to increase by an average of 6.63 % from the 1980-2018 average to the 2061-2099 average under RCP8.5, with all models showing a statistically significant increase. There is also an increase in average incident shortwave radiation, by 3.20 % over the same time period, with 22 models showing a statistically significant increase, 4 showing a statistically significant decrease and 4 showing no significant change. While these changes are modest, and, as we explain above, do not necessarily translate to equivalent changes at smaller scales, we do feel that, on reflection, we should and **will make the following changes in our revised manuscript: i) better-justify our reasoning for not using all of the climate variable data from the CMIP5 simulations; and ii) provide a more balance and up front consideration of the limitations of this approach (drawing on our analysis of the radiation projections) and the potential impacts on our estimates of twenty-first century glacier retreat.**

Please also provide more information on the resampling procedure.

Yes, we appreciate we could have provided more information on this. To be clear, we repeat the 1980-2018 WRF simulations i.e. we preserve the year-by-year sequencing of the climate variables. We preferred this approach rather than randomly resampling the data to preserve any multi-year cycling e.g. ENSO which may be present in the driving data. We also manually checked the data to ensure that this approach would not introduce any strange "jumps" in the driving data across the 2018 to 1980 crossover and we were satisfied that it would not. **We will provide additional information and justification on our resampling procedure in the revised manuscript.**

Have the authors performed a sensitivity study on the input data?

We've answered this question below under "Albedo modeling" where the same question is asked.

Albedo modeling: The authors mention a poor performance in the WRF albedo representations (lines 363-375). In particular, WRF largely overestimates the albedo. Hence, it is not surprising that the JULES-OGGM model shows more conservative mass balance projections for 2100 than other models (e.g., GloGEM, MAR2012, Figure C1).

This appears to be a source of confusion which is also picked up in one of the minor comments below. To be clear, the albedo simulations that we're referring to here are those from JULES, not WRF. We only take the incident radiation variables from WRF. **In the revised manuscript, we will clarify the text in this section to make it clear that we're talking about JULES, not WRF.**

However, the authors implemented many improvements to the glacier modeling procedures in comparison to previous studies (lines 524-528). Hence, I find it hard to estimate whether the more conservative estimates stem from the problematic albedo modeling and resampled model input data, or whether these are actually more realistic estimates for glaciers in this region. Have the authors performed sensitivity tests?

We appreciate this very thoughtful line of enquiry. Like them, we are also fascinated by the reasons for the more conservative estimates and whether these are more realistic than estimates from other models. We're quite sure, though, that the differences are not simply a result of the problematic albedo modelling and our approach to resampling the other climate variables. The referee may be aware of a recent study by Aguayo et al. (preprint) who compare the sources of glacier projection uncertainty in the Patagonian Andes using OGGM. They show that the historical climate data used to calibrate the model is actually a more important source of projection uncertainty than the future climate. They also show the significance of other aspects such as ice thickness and initial glacier coverage. This study suggests that these additional differences between our model setup and those of GloGEM and MAR2012 are also likely to be important and we allude to this in the discussion (lines 525-531):

"These differences likely stem from a range of sources in addition to differences in the glacier models themselves. For example, we used statistically-downscaled and bias-corrected climate projections rather than the raw CMIP 5 data. We also used different initial glacier coverage and thickness estimates based on the more refined maps of Drenkhan et al. (2018). Perhaps most-significantly, we tuned JULES-OGGM to geodetic data, while the GlacierMIP models were tuned to World Glacier Monitoring Service (WGMS) data. Indeed, this could explain why, even though the model evaluation showed JULES-OGGM to overestimate mass loss by a factor of two, it still shows to be more conservative than the GlacierMIP models."

We have consciously remained ambiguous on the exact sources of these differences as we appreciate that the reasons are likely to be multifaceted and a decomposition of these sources (e.g. through sensitivity analysis) is beyond the scope of what is already a comprehensive study. We don't see this as a negative of the paper, but rather something to inspire others to delve deeper into these unknowns. **However, we do feel that we could do much more in the discussion and conclusions to emphasise this as an important focal point for future work. We will do this in our revised manuscript.**

Minor comments:

Abstract: Please mention the grid spacing of your model in the abstract.

An interesting request! This is tricky as the model doesn't strictly have a grid spacing. We hope that this is clear from our explanation of the model setup in section 2.4.3 and Figure 3 where we note that:

"*Each glacier is represented by N ice-covered JULES grid boxes*".

These do not represent a strict "space", but rather the surface energy and mass balance at an elevation. We make this clear by stating that:

"The outputs from JULES, therefore, indicate how surface mass balance changes as a function of elevation only on each glacier."

These are used to downscale to the scale of the one dimensional OGGM flowline nodes, which themselves are spaced differently for each glacier:

"For each OGGM node of that glacier, the mass balance at the OGGM node elevation is extracted from the JULES grid box simulations (yellow dash arrows linking, Figure 3b and d), using linear interpolation to downscale to the resolution of OGGM."

It is therefore not possible for us to quote a particular grid spacing for JULES-OGGM.

Page 8 line 148-151: Do Dussaillant et al. (2019) provide yearly or seasonal data? Please specify.

They actually provide the data aggregated over a number of years. **We'll make this clear in the revised manuscript.**

Page 8 lines 165-170: Can the authors provide details on how the turbulent and latent heat fluxes and the ground heat flux were calculated?

We appreciate that we could provide more detail on the individual model fluxes, but we purposely decided not to do this, in part for conciseness and in part because the full list of equations are available in Best et al. (2011) which we refer the reader to in lines 155-156. We would prefer not to include these additional details as the manuscript if already quite comprehensive. Instead, we **propose to explicitly refer the reader to the Best et al. (2011) study in the descriptive text of the model energy balance (lines 166-170).**

Page 5 lines 121-122: "Grid spacing" and "resolution" refer to two different length scales and should not be used interchangeably (e.g., Grasso, 2000; Stull, 2015). It would be more appropriate to use "grid spacing" here and for similar cases.

Agree, **we will remove all references to "resolution" where they are not appropriate.**

Page 6 section 2.4.1: Can the authors please add some details on the start dates for summer and winter periods in the model?

We're a bit confused by this comment. What do you mean by "in the model?" You can't specify a start date for a season in the model. And how does this relate to the text on page 6 or in section 2.4.1? Could the referee please clarify this for us.

Page 11 lines 225-226: How did the authors come up with 10 grid box nodes with an equal spacing of 167 m elevation for an elevation difference of 2500 m between zmin and zmax? Please explain.

Good spot! It should say 278 m not 167 m. **This number will be corrected in the revised manuscript.**

Page 11 lines 236-238: Please provide details on the adjustments that were used for the shortwave radiation.

Yes, on reflection we could have provided more detail here on what we mean by "adjusted". To be clear, we're accounting for the fact that the WRF incident shortwave radiation flux is for a flat horizontal surface whereas the glaciers have a tilted surface. The amount of direct shortwave radiation (i.e. solar radiation received from the sun without having been scattered by the atmosphere) received on a tilted surface will be different to a flat surface. The ratio of the direct incident shortwave radiation received on a tilted surface with respect to a horizontal surface can be calculated from the geometric factor, $R_b = \cos\theta / \cos\theta_z$, where θ is the angle of incidence (i.e. the angle of the solar beam relative to the tilted glacier surface), and θ_z is the solar zenith angle. We calculated hourly R_b over the simulation period using the known glacier slope and aspect (taken from the SRTM digital elevation model) and hourly simulations of the position of the sun in the sky which was calculated using the Pysolar python package. This allowed us to scale the direct incident shortwave radiation dynamically based on the position of the sun in the sky and individual glacier geometries. Note that the WRF incident shortwave radiation flux includes both direct and diffuse (scattered) shortwave radiation. By default, JULES assumes that half of the incident shortwave radiation is direct. As we outline in the manuscript, we use the default parameterisation of JULES except where we have explicitly stated otherwise. There was no clear justification from straying from this for this study. Accordingly, the adjustment was only applied to half of the shortwave radiation flux provided by WRF. **We will include this additional detail in the revised manuscript.**

Page 11 line 240: typo

Agreed. **We'll address this in the revised manuscript.**

Page 11 lines 247-248: Have the authors used 10 JULES grid boxes for every glacier or used less grid boxes for glaciers that don't span the whole range of zmax-zmin?

Yes. **We'll clarify this in the revised manuscript.**

Page 12 lines 301-303: Are the results (section 3) based on one set of parameters for all glaciers, or was one set of parameters chosen per subregion (R1-R10)? Please specify.

The latter. We do feel that we have already made this clear in the manuscript and so propose not to change the text here. Lines 318-319 state that:

"*this regional approach was adopted for the remainder of the study*"

Page 16 lines 356-362: Have these two glaciers been part of the 30 glaciers used for calibration?

Yes. **We'll make this clear in the revised manuscript.**

Page 17 pages 363-375: WRF albedo modeling: The authors observed that the WRF-modeled albedo rarely falls below 0.8, but the observed albedo falls as low as 0.2 by the end of the dry season. The authors are using the WRF setup from Potter et al. (2023), who are using the Noah-MP land surface model. The default albedo parameterizations for land ice in Noah-MP are set relatively high and might need to be lowered (variable ALBICE in phys/module_sf_noahmp_glacier.F) for a value more consistent with bare ice in the study region. Have the authors explored changes in the WRF land surface model for a more realistic representation of albedo?

We address this in our response above, but to reiterate, we are not referring to the WRF albedo, only the JULES albedo.

Page 21 Figure 9 (e) and (f): Please specify which year you are referring to here (2020?)

Yes, we should have been clearer about this. We're using our initial area data (year 1998) derived by Drenkhan et al. (2018). **We'll make this clear in the revised manuscript.**

Page 23 Figure 10 and page 33 Figure E1: These figures are hard to read. Please increase the text and line sizes.

We agree, **we'll increase the text size for this figure in the revised manuscript.**

Page 23 lines 469-470: Can you provide an error estimate of the geodetic validation data used in this study?

Unfortunately, not. Although the geodetic dataset developed by Dussailant et al. (2018) includes error estimates, these are provided on a glacier-by-glacier basis using the RGI outlines. We used different glacier outlines, which meant that we had to calculate the geodetic mass balance of each glacier ourselves from the 30 m gridded data. To calculate the uncertainty, we'd need the raw analysis data that the authors used at each grid cell. This would be considerable work for something which we feel is desirable, but not fundamental to our study.

Page 24 lines 484-486: Can the authors give a brief overview of the current snow albedo routine in JULES?

Yes, while we did include a very brief overview of the processes included in this routine (lines 171-173), we appreciate that we could expand this slightly to provide additional details relevant to the study. **We'll add this text in the revised manuscript.**

Page 24 lines 491-496: Getting the net radiation correct in glacier modeling is a (main) challenge beyond tropical glaciers. Are there any conclusions for glacier calibration that can be drawn which are specific to tropical glaciers?

We agree, this is a challenge beyond tropical glaciers, but is one that clearly stands out as fundamental in our study, and therefore warrants it's place in our discussion. We feel that we have been rightly cautious about the transferability of our findings to other regions outside of the tropics and/or generalising our findings across all tropical glaciers or drawing conclusions about tropic-specific characteristics, given the scope of the study (e.g. spatial coverage, validation data, modelling approaches etc). Indeed, in our conclusions we make this clear by stating that:

"*results are not necessarily indicative of all glacierised basins in the tropics*"

and that:

"*we believe there is much to be learned from applying a physically-based, globally-capable model like JULES-OGGM to other basins inside and outside of the tropics*"

Even so, we do draw conclusions that are certainly of interest for modelling tropical glaciers in similar high-altitude, semi-arid settings such as those around the significance (or not) of sublimation processes.

Page 25 lines 526-527: I believe it is important here to mention which variables have been downscaled (i.e., temperature and precipitation), and that the other variables have been resampled for 2019-2100.

We completely agree and, as we mentioned in our response to the major comments, **we will add additional consideration of these aspects of the model setup and their implications for the results that we have presented in the discussion.**

References:

Grasso, LD. (2000). The Differentiation between grid spacing and resolution and their application to numerical modeling. Bulletin of the American Meteorological Society. 81 (3). 579-580. 10.1175/1520-0477(2000)081<0579:CAA>2.3.CO;2.

Stull, R. B. (2015). Practical meteorology: An algebra-based survey of atmospheric science. Department of Earth, Ocean & Atmospheric Sciences, University of British Columbia, Vancouver, BC. https://doi.org/10.14288/1.0300441.

References:

Aguayo et al., preprint, https://doi.org/10.5194/egusphere-2023-2325

Best et al., 2011, https://doi.org/10.5194/gmd-4-677-2011

Drenkhan et al., 2018, https://doi.org/10.1016/j.gloplacha.2018.07.005

Dussaillant et al., 2019, https://doi.org/10.1038/s41561-019-0432-5

Potter et al., 2023, https://doi.org/10.1038/s41612-023-00409-z