***Interactive Comment on* "Exploring the ability of the variable-resolution CESM to simulate cryospheric-hydrological variables in High Mountain Asia*" by* René R. Wijngaard et al.**

We gratefully acknowledge the reviewer for his/her remarks and suggestions, which improved the quality of the manuscript significantly. We have carefully considered the suggestions of the reviewer and we provide a point-by-point response to the reviewer's comments. For clarity, the reviewer's comments are given in bold italics and the responses are given in plain text. References that do not refer to those in the main manuscript are listed below. The manuscript will be modified according to the responses that are given to the comments.

Summary

Wijngaard et al. in "Exploring the ability of the variable-resolution CESM to simulate cryospheric-hydrological variables in High Mountain Asia" evaluate a first-of-its-kind study on High Mountain Asia glaciers using the variable-resolution capabilities in the Community Earth System Model at 9 km horizontal refinement. The authors do an admirable job of comprehensibly evaluating all aspects of snow energy and mass balance (SEB and SMB) drivers of HMA mountain glaciers. Unfortunately, it is identified that over the 1979-1998 period HMA mountain glacier SMB loss in VR-CESM is 10x higher than expected (expected SMB is -20 GT/yr, VR-CESM SMB is, at best, -200 GT/yr), although not due to any fault of the authors. The authors provide several plausible bias sources and hypotheses to test in future VR-CESM studies.

Overall, I think the paper fits perfectly within the scope of The Cryosphere and could be, given more work, an extremely valuable contribution to the scientific community. The findings have both scientific and societal impact as mountain glacier modeling, particularly in Earth system models, is a recent model development across the community of models (requiring comprehensive historical evaluation studies such as this one) and glacial melt supplements water supplies and poses significant flood hazards (e.g., glacier lake outburst flood events) to billions of people that reside downstream of them.

I think there are still several major(ish) revisions that need to happen prior to this paper being accepted. Most of my suggested revisions are minor, however a few (if feasible) may require some time to address, and I wouldn't want the authors to be time pressured by a quick turnaround with a suggestion of minor revisions.

We thank the reviewer for his/her evaluation of the manuscript. We have tried to address all concerns.

Suggested Revisions
Line 14 – change "could help…" to "can help through targeted grid refinement"

We have changed "could help…" to "can help through targeted grid refinement…".

Line 18 – delete "to evaluate the outcomes"

We have deleted "to evaluate the outcomes".

Line 22 – add "but is still underestimated"

We have added "but is still underestimated".

Line 37 – change "differs per region" to "is regionally dependent"

We have changed "differs per region" to "is regionally dependent".

Line 41 – change "as a response to climate change" to "in response to climate change"

We have changed "as a response to climate change" to "in response to climate change".

Line 52-53 – consider citing - Li, D., Lu, X., Walling, D.E. et al. High Mountain Asia hydropower systems threatened by climate-driven landscape instability. Nat. Geosci. 15, 520–530 (2022). https://doi.org/10.1038/s41561-022-00953-y

Thanks for the suggestion. We will include the citation of Li et al. (2022) in the manuscript.

Line 60 – change "with or without" to "with"

We have changed "with or without" to "with".

Line 63 – change "high horizontal…" to "fine horizontal…across many glaciers"

We have changed "high horizontal…for many glaciers" to "fine horizontal…across many glaciers".

Line 71 – change "giving" to "providing"

We have changed "giving to "providing".

Line 79 and Line 99 – consider citing - Rhoades, A. M., Ullrich, P. A., Zarzycki, C. M., Johansen, H., Margulis, S. A., Morrison, H., et al. (2018). Sensitivity of mountain hydroclimate simulations in variable-resolution CESM to microphysics and horizontal resolution. Journal of Advances in Modeling Earth Systems, 10, 1357– 1380. https://doi.org/10.1029/2018MS001326

Thanks for the valuable suggestion. We will include the citation of (Rhoades et al., 2018) in the manuscript.

Line 94 – FWIW Rhoades et al. 2018 provides model throughput timing for a wider range of refinement regions (including a 7km refinement patch)

Thank you for the information. It is interesting to see that for about the same number of processors, the HMA_VR7 grid can produce about 35% (~0.7 SYPD) of the number of simulated years that the CAL_VR7 grid can produce within one wall-clock day (1.96 SYPD). This is mainly related to the larger domain that is covered by the 7 km refined patch in the HMA_VR7 simulations.

Line 102 – change "this far" to "thus far"

We have changed "this far" to "thus far".

Line 106 – why was 1979-1998 chosen rather than a period that encapsulates more of the satellite record?

Most AMIP-style runs in CESM start by default in 1979 and cover the period 1979–1998, which is the reason we also followed this approach. However, we agree with the reviewer that a later period that encapsulates more of the satellite record is more beneficial and helpful for evaluating the model outcomes. We expect that future model simulations will also cover later period, enabling comparisons with satellite products, such as those from the Sentinel project.

Line 152 – is a constant lapse rate and relative humidity with elevation an appropriate assumption? At the very least, is there a study to cite here?

The constant relative humidity with elevation in CLM is needed to control the amount of water vapor with elevation, i.e., to prevent conservation errors in CESM. Both relative and specific humidity are computed in CAM and then passed to CLM. Here specific humidity is downscaled based on the assumption of constant relative humidity, downscaled air temperature and downscaled saturation water vapor pressure, where downscaled saturation water vapor pressure is scaled by a polynomial fit between air temperature and saturation vapor pressure (Flatau et al., 1992). Since the temperature generally decreases with elevation, the saturation vapor pressure and specific humidity also decrease with altitude. By keeping the relative humidity constant, the actual water vapor pressure is scaled down in a similar manner, which prevents the model from simulating a surplus or deficit in actual water vapor. Based on the idea that a constant relative humidity prevents conservation errors, we think it is an appropriate assumption. We will add a citation to the study of Lipscomb et al. (2013).

Line 156 – cite the topography dataset developer manuscript/data archive.

We will include a citation to a data archive on Zenodo (https://doi.org/10.5281/zenodo.7864689) to the manuscript.

Equation 1 and 2 – can black/brown carbon deposition influence SEB and SMB in VRCESM? If so, is black/brown carbon deposition too high and helping to drive the SMB bias? Also, can surface melt pools and flow channels develop (influences albedo and, potentially, energy transport through glacier)?

Thank you for raising these points. Our CESM2 simulations are forced with both prescribed CMIP6-based tropospheric and stratospheric aerosols that include mineral dust, sea salt, and black carbon (BC), amongst others. Aerosols can potentially influence SMB and SEB in various ways in CESM2. These include a warming of the atmosphere due to aerosol-radiation interactions and reductions in snow albedo due to darkening of the snow surface. The aerosol-radiation interactions in CESM are simulated by the Modal Aerosol Model with 4 modes (MAM4; Liu et al., 2016) combined with the Rapid Radiative Transfer Model for GCMs (RRTMG; (Iacono et al., 2008)). Aerosol-induced changes in snow albedo and radiative heating transfers in the snowpack are simulated by SNICAR (Flanner et al., 2007), the radiative transfer model coupled to CLM.

Aerosol-radiation interactions and effects of aerosol deposition, such as black carbon (BC) and dust deposition, on snow darkening in CESM have previously been investigated by Rahimi et al. (2019), amongst others. The authors investigated the effects of BC and dust deposition on the South Asian summer monsoon and snow darkening with a 14 km regionally refined grid in combination with CESM1.2. According to the authors, BC deposition mainly result in an intensification of summer monsoon rains via aerosol-radiation interactions. Also, BC deposition has a snow-darkening effect that leads to reductions in SWE, particularly over the western Tibetan Plateau (corresponding to SW-HMA in our study). The authors show, however, that the effect of BC deposition on snow darkening is weaker in the VR simulations than in the 1-degree simulations due to the lower amount of BC deposition that is simulated by the VR grid.

To have a general understanding of the BC concentrations simulated by the NE30 and HMA VR grids, we looked at the BC concentrations in the top snow layer, defined as the ratio between the mass of BC and the mass of snow in the top snow layer (the thickness of the top snow layer is 1–3 cm, depending on the snow depth and related amount of snow layers). Another commonly used method is to define the BC concentrations as the ratio between the total BC deposition from the atmosphere and the total precipitation (e.g., He et al., 2014). However, this method introduces uncertainties in the calculation of in-snow BC concentrations, which is mainly related to the use of total precipitation that can overestimate the amount of snowfall and BC removed by snow and is prone to biases relative to observed precipitation (He et al., 2014). Therefore, we did not choose for this method and decided to focus on the BC concentrations in the top snow layer.

Tables RC2.1 and RC2.2 show the spatial mean and standard deviation of seasonal BC concentrations in the top snow layer for NE30 and HMA_V7b simulation outputs, respectively. The highest BC concentrations are simulated during spring and summer, which is most likely caused by the aging of snow and reduced influx of fresh snow (due to higher temperatures). Spatially, the higher BC concentrations are simulated in SE-HMA, which is more prone to BC deposition originating from the Indo-Gangetic Plains (not shown). Looking at the differences between the simulated NE30 and HMA_VR7b BC concentrations, we observe that the NE30 grid simulates higher BC concentrations than the VR grid, which is also in line with the findings of Rahimi et al. (2019). Compared to observations, such as those listed in He et al. (2014) or Kang et al. (2020), the simulated BC concentrations in the top snow layer fall in range of what is observed over HMA. However, the range of observed values is large (i.e., between 16 ng/g and 3000 ng/g; Kang et al., 2020). Also, most observations are based on measurements conducted over a short timeframe (e.g., one season). Nonetheless, based on the available observations we think that the BC deposition is not too high, and that the impact of the BC deposition on the SMB bias is relatively small.

Table RC2.1. Mean seasonal BC concentrations (ng/g) in top snow layer as simulated by the NE30 grid in the four HMA subregions (Figure 1b). The numbers between brackets denote the standard deviation of the BC concentrations. The BC concentrations are derived for grid cells where glaciers are present, and the mass of snow is higher or equal to 0.3 kg/m^2 (i.e., equivalent to 1 mm of snow, assuming a snow density of 300 kg/m^3). The latter was necessary to filter out unrealistically high BC concentrations due to the limited amount of snow in the NE30 simulations during summer and autumn.

	NW-HMA	NE-HMA	SW-HMA	SE-HMA
DJF	84 (50)	51 (31)	76 (45)	109 (39)
MAM	136 (89)	75 (44)	125 (96)	188 (119)[1]
JJA	162 (110)	75 (117)	386 (267)	485 (434)[1]
SON	54 (20)	38 (24)	128 (84)[1]	90 (68)

[1] To filter out unrealistically high BC concentrations, grid cells where the mass of BC is higher than the global mean (~0.03 kg/m^2) are filtered out.

Table RC2.2. Same as Table RC2.1 but for HMA_VR7b.

	NW-HMA	NE-HMA	SW-HMA	SE-HMA
DJF	42 (26)	56 (47)	73 (64)	144 (72)
MAM	85 (48)	74 (51)	108 (76)	208 (90)
JJA	93 (101)	32 (65)	238 (241)	184 (224)
SON	62 (29)	41 (31)	97 (70)	92 (46)

As for the surface melt pools and flow channels, these cannot develop in CLM, since CLM does not take the complex topography of mountain glaciers into account. To sufficiently simulate surface melt pools, flow channels, and other important glacier features (e.g., crevasses or ice cliffs), models are required with a high horizontal resolution up to 1 m (e.g., Bonekamp et al., 2020).

Line 195 – delete "more"

We have deleted "more".

Line 218-219 – could the authors add the observed max elevation for each HMA region so that readers can contrast with model representation max elevations?

We will include the following table, which shows the observed maximum altitude for each HMA subregion:

Table 1. Maximum altitude per HMA subregion (Figure 1b) as observed for NE30, NE120, HMA_VR7 and GMTED2010.

	NW-HMA	NE-HMA	SW-HMA	SE-HMA
NE30	5162 m	5170 m	4978 m	4452 m
NE120	5684 m	5526 m	5513 m	5369 m
HMA_VR7	6167 m	5778 m	5870 m	6228 m
GMTED2010	7325 m	7099 m	8190 m	8625 m

Also, we have changed the following sentence (L218-219):
"Over HMA, the maximum altitude increases from 5170 m and 5684 m for NE30 and NE120, respectively, to 6227 m for HMA_VR7."

into

"Over HMA, the maximum altitude increases from 5170 m and 5684 m for NE30 and NE120, respectively, to 6228 m for HMA_VR7. These maxima are observed in the northeastern, northwestern, and southeastern HMA subregions (Figure 1b), respectively (Table 1)."

Line 224 – glaciers are assumed constant? Is this true for these simulations or just the default setting? If so, how would keeping glaciers constant (I'm guessing areal extent of glacier cover in grid cell?) shape the results?

Yes, the areal extent of glaciers is assumed to be constant throughout the simulation periods. To simulate dynamical changes in glacial extent, CLM needs to be coupled with the Community Ice Sheet Model (CISM), the dynamic ice sheet component of CESM. In the CESM model version we used, active CLM-CISM coupling is only possible over the Greenland Ice sheet. Ongoing developments, however, will enable active CLM-CISM coupling over mountain regions such as HMA. Therefore, we expect that in future CESM model versions, it will be possible to simulate dynamical changes in HMA glaciers.

The effects of constant glacier extent on simulation outcomes depend on the amount of time a glacier needs to respond to climatic changes. Small glaciers have generally a short response time and are more vulnerable to climate change than large glaciers with a long response time. For instance, (Shea et al., 2015) found that small glaciers in the Everest region (i.e., located in the SE-HMA subregion) have a response time of 20–50 years, whereas large debris-covered glaciers have a response time of 200–500 years. Since our simulations cover a 20-yr period (1979–1998), we expect that the effects of constant glacial extent on simulation outcomes are rather small. However, to quantify these effects, an active coupling between CLM and CISM would be required.

Line 235 and 250 – could these fixed, solely temperature-based thresholds influence SMB bias? Would implementing Jennings et al. 2018 temperature-relative humidity-based rainsnow thresholds help (see Equations 3-4 in Jennings et al. 2018) to increase the probability of more snowfall and a more positive SMB (see discussion on hydrometeor energy balance in Jennings et al. 2018 for physical intuition on why accounting for humidity matters for snowfall)? This might be difficult given computational limitations (although several authors are at NCAR and may have access to additional computational time on Cheyenne), but could the authors run an experiment (5-10 years) with the HMA 7km grid and simply swap in the Jennings et al., 2018 temperature-humidity based rainsnow partition scheme to compare how SMB biases are altered (maybe add to Table 3)?

Jennings, K.S., Winchell, T.S., Livneh, B. et al. Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. Nat Commun 9, 1148 (2018). https://doi.org/10.1038/s41467-018-03629-7

Thank you for raising this point. The Jennings et al.-based temperature-based thresholds that are used in the HMA_VR7b simulation have shown an increase in the amount of snowfall relative to the temperature-based thresholds used in the HMA_VR7a simulation. Figure 2 of Jennings et al. (2018) plots the relative humidity (RH) and surface pressure (PS) against surface air temperature (TS). The figure shows that, for example, for Ts = 2 °C, lower RH and PS values result in a higher snowfall frequency. In this light, using a TS-RH-based threshold (Jennings Equation 3) or a TS-RH-PS-based threshold (Jennings Equation 4) could potentially help to increase the snowfall frequency and lower the SMB bias. Jennings Figure 6 shows, however, that on average the Eq. 3 and Eq. 4 based methods result in a lower snowfall frequency than the 2°C 50% rain-snow TS threshold used for the HMA_VR7b simulation. Nonetheless, the lower snowfall frequency probably does not occur over HMA as shown in Jennings Supplementary Figure 2, which shows a positive difference over HMA between the Eq. 4 and Eq. 3 based methods. This could mean that compared to the fixed TS-based threshold used in this study, the Eq. 4 based method also results in more snowfall over HMA. Therefore, it could be worth to exploring the effects of the Eq. 4 based method on the SMB bias.

Unfortunately, the computational resources are currently limited and mainly reserved for nudging experiments over HMA. For this reason, implementing Jennings Eq. 4 is not possible in the short term. However, we would like to consider implementing Jennings Eq. 4 and exploring its effects in future simulations when allocation will be used to improve land-atmosphere coupling. Also, we will add some text to the discussion in the manuscript to point to the limitations of using a fixed TS-based threshold.

Line 138, 232 and 247 – the assumption of capping snow depths at 1 m and 5 m w.e. in HMA seems like another culprit for SMB bias. How often does snow depth hit the cap over the 20-year simulations? Could the authors provide a cumulative annual "snow loss" estimate due to snow capping and compare contrast with SMB bias (I'm guessing this might influence refreezing portion of SMB in Table 3 values)?

The snow depth generally hits the cap only in those grid cells where a positive SMB is simulated. These grid cells are located in the southernmost part of HMA. The cumulative annual "snow loss" due to snow capping, which can be expressed as the excess amount of snow due to snow capping, amounts to about 2 (+/- 1) Gt yr^{-1} for HMA_VR7a and about 1 (+/- 0.1) Gt yr^{-1} for HMA_VR7b. This means the increased maximum allowed snow depth in HMA_VR7b has a positive impact on the cumulative annual "snow loss" and has resulted in increased refreezing (from 29 Gt yr^{-1} for HMA_VR7a to 32 Gt yr^{-1} for HMA_VR7b), which itself has a positive impact on the SMB bias. Also, the cumulative annual "snow loss" cannot be compared with the SMB bias, which is more than a factor 100 larger than the "snow loss". In other regions, the snow depth does not reach the cap and "snow loss" due to capping cannot occur. From the perspective of CLM, the SMB is zero in these grid cells where snow is present or negative when snow is absent. In these regions, we simulate a negative SMB, which can be attributed mainly to large amounts of ice melt (and thus the absence of snow). Therefore, we think that the assumption of capping snow depths cannot be a culprit for the SMB bias.

Line 265-290 – the authors might consider evaluating SWE model estimates from VRCESM compared with Liu et al. 2021 (note the period of record does not overlap with the VR-CESM simulations, but a climatological comparison might still be useful). In addition, ERA5-Land might also be useful too. This would be especially insightful for the snow capping assumption/issue. "…It can be accessed through https://nsidc.org/data/HMA_SR_D/ (last access: 22 April 2021) or Https://doi.org/10.5067/HNAUGJQXSCVU (Liu et al., 2021). The dataset is provided as NetCDF files for each 1∘×1∘ tile shown in Fig. 1, available at 16 arcsec (~500 m) and daily resolution from WYs 2000 to 2017…"

Liu, Y., Fang, Y., and Margulis, S. A.: Spatiotemporal distribution of seasonal snow water equivalent in High Mountain Asia from an 18-year Landsat–MODIS era snow reanalysis dataset, The Cryosphere, 15, 5261–5280, https://doi.org/10.5194/tc-15-5261-2021,2021.

https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview

Thank you for your suggestions. We decided not to use ERA5-Land and the datasets of Liu et al. (2021) for several reasons. The datasets of Liu et al. (2021) are very valuable, but require a lot of storage and are computationally heavy. For example, 5 years of output for one 1-degree tile (e.g., 28-29N; 84-85E) require around 3 GB of data. ERA5-Land and ERA5 are both known to overestimate snow depth and snowfall over HMA (Orsolini et al., 2018). For this reason, we initially selected JRA-55 as the evaluation dataset since it has the best performance in the representation of snow depth. We have, however, found an alternative dataset that also can be used for snow depth evaluation, which is the High Asia Refined analysis version 2 (HARv2; Wang et al., 2021). This dataset is based on WRF simulations that are dynamically downscaled with ERA5 and corrected with snow depth from JRA-55. Also, the spatial resolution of HARv2 (~10 km) and period (1980–1998) match good with the HMA VR simulations. Figure RC2.1 shows the snow depth differences between the NE30 and HMA VR configurations, and the reanalysis-based HARv2 dataset. Snow depth is mainly overestimated over SW-HMA and SE-HMA during winter and spring, in part because of the cold bias. During autumn (over SW-HMA), and winter (over NW-HMA) the snow depth is slightly underestimated, which interestingly matches better with the negative snow cover biases during autumn and winter in the respective regions (Figure 9a). In the manuscript we will replace the JRA-55 based snow depth evaluation by the HARv2 snow depth evaluation.

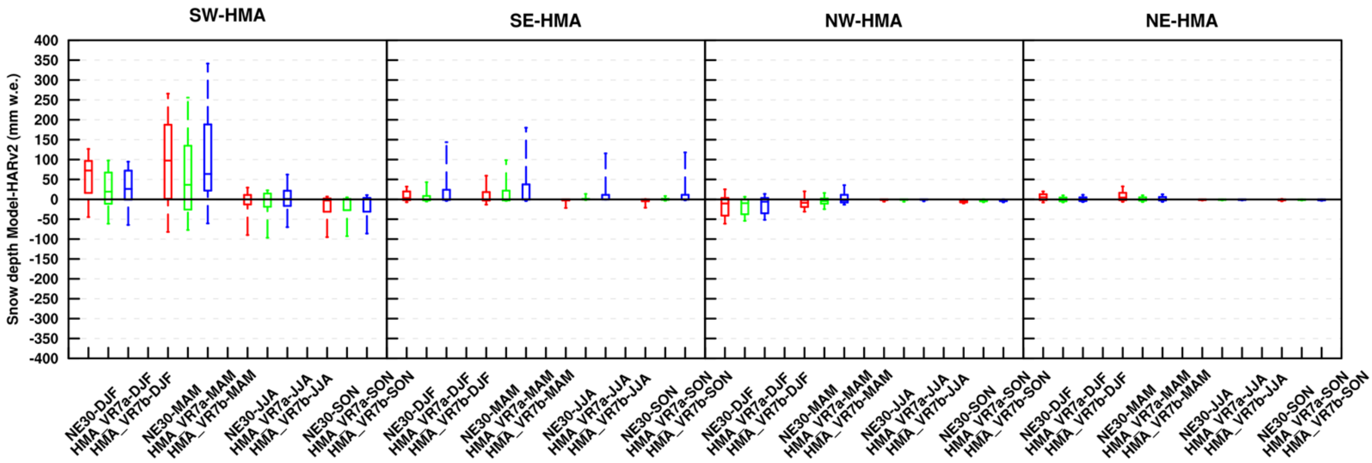

Figure RC2.1 Boxplots of snow depth differences (mm w.e.) between the simulation outputs of NE30 (red), HMA_VR7a (green), HMA_VR7b (blue), and HARv2 for each season and HMA subregion (shown in Figure 1b). The box represents the biases between the 25th and 75th percentile, the line in the box denotes the median, and the whiskers represent the 10th and 90th percentile of snow depth differences.

Line 305 – eddy or geopotential height?

It is eddy geopotential thickness, which is defined as the deviation of the geopotential thickness from the zonal mean of the geopotential thickness, where the geopotential thickness is calculated as the difference between geopotential heights at 500 hPa and 1000 hPa.

Line 314-321 – very informative! It might be worth mentioning here that 2m surface temperatures don't always correspond/agree with free/upper atmosphere warm biases, particularly over HMA (the authors later show that 2m surface temperatures are colder than expected in most HMA regions) to hedge reader expectation/confusion.

We have changed the following sentence (Line 314-315) from:

"The upper tropospheric summer temperature biases (relative to ERA5) can help to understand the effects of atmospheric biases on temperature-sensitive surface variables such as ice melt and snow melt."

to

"The upper tropospheric summer temperature biases (relative to ERA5) can help to understand the effects of atmospheric biases on temperature-sensitive surface variables such as ice melt and snow melt. However, atmospheric biases do not necessarily need to correspond with surface temperature biases as we will show in Section 3.3."

Figure 3 – for the temperature anomaly plots, the authors might consider decreasing the number of color bins to 0.5 deg C bins (current precision is hard to interpret).

We will decrease the number of color bins to 0.5 deg C as suggested by the reviewer.

Figure 6 – could the authors add subpanel labels (a, b, …) and point to them while describing physical interpretation in Lines 343-356 (which is well explained, but pointing to specific plots might help those that aren't super familiar with monsoonal/precipitation dynamics in this region). Define 2nd y-axes on first column of subpanels (or remove).

We will add subpanel labels and point to them accordingly in the manuscript. Also, we will add a label to define the 2nd y-axis.

Line 358-376 – should median biases be stated? I think the authors should at least highlight that most of the distribution of the box-and-whiskers fall below the 0 deg C bias line (cold bias), save for JJA in the NW- and NE-HMA, and particularly VR-based distributions.

We think it is good to quantify the biases to give an impression on the magnitude of the biases for the reader. In our opinion, the median of the biases could be a good reference point. For this reason, we prefer to keep the median biases in the text.

Figure 7 – "absolute temperature" shouldn't all values then be positive?

Our aim was to distinguish between "absolute" and "relative" changes, where "absolute" change is expressed in terms of a quantitative change and a "relative" change is expressed in terms of a percent change. To avoid further confusion, we will remove "absolute" from the manuscript text and figures' captions.

Figure 8 – the use of a-b and c-d are incorrect, should be a-c (rainfall) and b-d (snowfall)

We have changed "for rainfall (mm month-1) (a-b) and snowfall (mm month-1) (c-d)" to "for rainfall (mm month-1) (a-c) and snowfall (mm month-1) (b-d)"

Line 368-376, 323-330 and 358-376 – Bambach et al. 2022 and Rhoades et al. 2018 (and others) have shown that a cold bias with elevation is also seen in non-glaciated regions in mid-latitude mountain regions (and without the application of the downscaling/ECs). Do the authors think that the thinner clouds could be a culprit for the colder surface temperatures (Figure 7)? I'm thinking during the day there could be more shortwave insolation, but in the night (lower cloud fraction/thinner clouds) there could be enhanced radiative cooling with the net daily balance being negative? Or is this surface temperature cold bias driven more by a lack of longwave feedbacks and/or minimal boundary layer turbulence over snow? Slater et al. (2001) had a hypothesis for a positive feedback loop that creates stable boundary layers, particularly over winter/snow conditions (see Figure 8 in Slater et al., 2001). Also, if this cold bias is fixed, it would likely worsen the SMB bias (unless the cold bias is partly due to the melt energy extracted from the atmosphere to the glacier, as the authors hypothesize).

Bambach, N. E., Rhoades, A. M., Hatchett, B. J., Jones, A. D., Ullrich, P. A., & Zarzycki, C. M. (2022). Projecting climate change in South America using variable-resolution Community Earth System Model: An application to Chile. International Journal of Climatology, 42(4), 2514– 2542. https://doi.org/10.1002/joc.7379

Rhoades, A. M., Ullrich, P. A., Zarzycki, C. M., Johansen, H., Margulis, S. A., Morrison, H., et al. (2018). Sensitivity of mountain hydroclimate simulations in variable-resolution CESM to microphysics and horizontal resolution. Journal of Advances in Modeling Earth Systems, 10, 1357– 1380. https://doi.org/10.1029/2018MS001326

Slater, A. G., Schlosser, C. A., Desborough, C. E., Pitman, A. J., Henderson-Sellers, A., Robock, A., Vinnikov, K. Y., Entin, J., Mitchell, K., Chen, F., Boone, A., Etchevers, P., Habets, F., Noilhan, J., Braden, H., Cox, P. M., de Rosnay, P., Dickinson, R. E., Yang, Z., Dai, Y., Zeng, Q., Duan, Q., Koren, V., Schaake, S., Gedney, N., Gusev, Y. M., Nasonova, O. N., Kim, J., Kowalczyk, E. A., Shmakin, A. B., Smirnova, T. G., Verseghy, D., Wetzel, P., & Xue, Y. (2001). The Representation of Snow in Land Surface Schemes: Results from PILPS 2(d), Journal of Hydrometeorology, 2(1), 7-25. Retrieved Feb 13, 2023, from https: //journals.ametsoc.org/view/journals/hydr/2/1/1525-7541_2001_002_0007_trosil_2_0_co_2.xml

There are several potential explanations for the cold biases in the NE30 and HMA VR simulation outputs besides the combination of better resolved topography and EC downscaling we mentioned in the manuscript. First, we consider it likely that the cold bias is partly a result of uncertainties in the observation/reanalysis-based WFDEI we used to evaluate surface temperatures. Since WFDEI is bias-corrected using gridded observations of GPCC, the accuracy of the data relies on the availability of meteorological measurements that are scarce in High Mountain Asia, especially at higher altitudes and in the more remote domains of HMA. The lack of measurements could result in temperature overestimates, but also in precipitation underestimates (Palazzi et al., 2015; Immerzeel et al., 2015; Gu et al., 2012; Lalande et al., 2021), which can to some extent explain the cold and wet biases that are visible in the NE30 and HMA VR simulation outputs.

Second, the thinner clouds can contribute to the colder surface temperatures as suggested by the reviewer, especially during wintertime when the net radiative balance is already negative (Table RC2.3). During the day, the thinner clouds could result in more shortwave insolation, which in summer can contribute to more melting. During the night, when the radiative balance is negative, the thinner clouds could result in enhanced radiative cooling. The enhanced radiative cooling could then eventually lead to a negative net daily radiative balance, especially during wintertime. Therefore, the thinner clouds could also explain to some extent the cold biases that are present in the simulation outputs.

Third, the cold biases could partly be a result of the stability-induced cooling feedback (Figure 8 in Slater et al., 2001). During winter, the net radiative balance is negative (Table RC2.3) and sensible/latent heat fluxes are close to zero. Further, the ground heat flux is positive towards the surface, partly countering the negative net radiative balance. Radiative cooling occurs, which causes the surface temperature to be lower than the overlying air temperature (i.e., a decoupling of the surface from the atmosphere). Via the stability-induced cooling feedback, the radiative cooling can be enhanced, which eventually results in cold biases, particularly during wintertime. One way to prevent a stability-induced cooling feedback could be to improve the coupling between land and atmosphere as already outlined in the manuscript. We will add text discussing the possible links between cold biases and the thinner clouds, stability-induced cooling feedback, and the uncertainties in the observation/reanalysis-based WFDEI.

Table RC2.3. Mean wintertime (DJF) surface-energy-balance (SEB) fluxes (W m^{-2}) for NE30, HMA_VR7a (HMA1), HMA_VR7b (HMA2), JRA55 and CERES-EBAF (C-EBF). SEB fluxes are defined positive towards the surface and represent the mean fluxes in glaciated grid cells (encompassing all land units), averaged over the entire HMA region and the periods 1979–1998 (for NE30, HMA_VR7a, HMA_VR7b, and JRA55) and 2001–2020 (for CERES-EBAF).

SEB fluxes (W m^{-2})	NE30	HMA1	HMA2	JRA55	C-EBF
SW_d	162	168	164	162	133
SW_u	87	85	83	80	42
SW_{net}	75	83	80	81	92
LW_d	179	165	171	175	207
LW_u	261	253	256	263	277
LW_{net}	-82	-88	-85	89	-71
R_{net}	-7	-5	-5	-8	21
SHF	~0	-0.5	-2	7	-
LHF	-5	-4	-4	-14	-
GHF	13	12	12	16	-
MHF	2	2	2	2	-
Refreezing	0.7	0.6	0.6	-	-

Line 377-391 – could the authors provide cumulative annual/monthly precipitation totals to better contextualize monthly average biases? I'm not familiar with the HMA region's cumulative precipitation totals and can't contextualize the +/- monthly biases stated.

We will add a table (Table S2) to the Supplementary Information that contains the mean winter (DJF) and summer (JJA) 2m temperature, rainfall, snowfall, snow cover, and snow depth for NE30, HMA_VR7a, HMA_VR7b, and the observation/satellite/reanalysis-based datasets. The seasonal sums/means are calculated over the HMA subregions for the period 1979–1998.

Table S2. Mean winter (DJF) and summer (JJA) 2m temperature, rainfall, snowfall, snow cover, and snow depth for NE30, HMA_VR7a, HMA_VR7b, WFDEI (temperature + precipitation), NSIDC (snow cover), and HARv2 (snow depth). The seasonal means/sums are calculated over the HMA subregions (Figure 1b) for the period 1979–1998.

	SW-HMA		**SE-HMA**		**NW-HMA**		**NE-HMA**	
2m Temperature (°C)	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**
NE30	-12	13	-1	15	-9	18	-14	12
HMA_VR7a	-16	10	-2	14	-11	16	-14	12
HMA_VR7b	-15	10	-1	14	-11	17	-14	12
WFDEI	-8	12	0	15	-10	16	-12	10
Rainfall (mm)	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**
NE30	25	383	52	1109	14	84	1	342
HMA_VR7a	21	420	37	1171	11	82	0.4	290
HMA_VR7b	19	435	37	1130	11	84	0.3	281
WFDEI	48	145	38	740	16	75	1	163
Snowfall (mm)	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**
NE30	237	9	76	4	62	2	30	20
HMA_VR7a	193	33	58	41	50	12	17	34
HMA_VR7b	208	43	62	53	61	14	19	39
WFDEI	99	12	23	50	45	3	9	18
Snow Cover (%)	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**
NE30	81	2	30	0.1	66	0.2	61	0.2
HMA_VR7a	83	5	28	2	65	1	53	1
HMA_VR7b	83	7	28	2	67	1	52	1
NSIDC	75	42	24	9	68	17	25	2
Snow Depth (mm w.e.)	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**	**DJF**	**JJA**
NE30	157	27	26	0.6	42	2	22	0.2
HMA_VR7a	121	24	27	8	38	1	15	0.4
HMA_VR7b	130	39	50	28	43	2	16	1
HARv2	96	34	17	5	57	4	14	1

Line 381 – change "worst" to "worse"

We have changed "worst" to "worse".

Line 386 – I'm curious how accounting for both temperature and humidity within the Jennings et al. 2018 rain-snow partitioning scheme(s) would alter this snowfall bias

We also wonder how the Jennings rain-snow partitioning scheme would alter the snowfall bias. We hope to find out the performance of Jennings rain-snow partitioning scheme in future simulations.

Line 400-406 and Figure 9 – I'd recommend comparison with Liu et al. (2021) to better verify snow depth/SWE. The authors might also consider using ERA5-Land. ERA5-Land could be useful for other surface variable comparisons (esp. since it's at a much more comparable resolution, ~9km, to VR simulations and the authors wouldn't need to coarsen for comparison).

Liu, Y., Fang, Y., and Margulis, S. A.: Spatiotemporal distribution of seasonal snow water equivalent in High Mountain Asia from an 18-year Landsat–MODIS era snow reanalysis dataset, The Cryosphere, 15, 5261–5280, https://doi.org/10.5194/tc-15-5261-2021, 2021.

https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview

To verify snow depth, we have chosen to compare snow depth with the High Asia Refined analysis version 2 instead of JRA-55.

Line 411-413 – this finding matches one of my earlier comments/questions about LW feedbacks potentially driving surface temperature cold biases. Although fixing this bias might further impact SMB bias since snow/ice are nearly blackbodies in LW spectrum…

We think that improving the subgrid coupling between land and atmosphere by introducing CLM patch information into CAM could reduce the cold bias. Also, nudging the vertical levels of CAM with data from ERA5 could reduce the cold and SMB bias by increasing the cloud cover over HMA. We are planning to conduct nudging experiments and experiments with improved land-atmosphere coupling schemes (with CLASP, Waterman et al., 2022) that will hopefully improve the SMB and temperature biases.

Line 427 – I may have misinterpreted this, but I thought that there was reduced/thinner cloud cover (see Line 327-330) not greater, as stated here.

Yes, relative to NE30 the HMA VR simulations have reduced/thinner cloud cover, but relative to HMA_VR7a, HMA_VR7b has greater cloud cover due to the cloud tunings that were implemented in HMA_VR7b. We will clarify the text by making clear that the greater cloud cover mentioned here is an observation relative to the cloud cover simulated in HMA_VR7a.

Line 436 – the extensive ice melt (and heat extraction from environment) could also be driving the surface temperature bias?

Although the surface temperature cold bias is smaller during summer, we think it likely that the extensive ice melt can also contribute to the surface temperature bias. As mentioned in the manuscript, the extensive melt requires heat extraction from the environment, causing the ground heat flux to be positive towards the surface and the surface temperature to be lower. Over the ablation zones of the Greenland Ice Sheet, cold temperature biases are also present during summer that, to some extent, overlap with regions where ground heat fluxes are positive and melt heat fluxes are large (Figures 4 and 7 of Kampenhout et al., 2020). We will mention in the manuscript that the extensive ice melt via heat extraction from the environment likely also contributes to the cold temperature biases.

Line 441-457 and Figure 13 – given that SMB doesn't appear to be water limited (i.e., VRCESM produces too much precipitation/snowfall in most HMA regions in Figure 8), how much does the atmospheric variable SMB downscaling method influence the SMB bias (particularly temperature/surface energy fluxes)? I'm wondering how the use of a fixed lapse rate (6 K/km and 32 W/m^2*km) and constant relative humidity with elevation shapes the negative SMB, particularly in mountains where lapse rates can vary quite a lot, even ones much smaller and with less heterogeneity than HMA (see Lute and Abatzoglou, 2020)

Lute, AC, Abatzoglou, JT. Best practices for estimating near-surface air temperature lapse rates. Int J Climatol. 2021; 41 (Suppl. 1): E110– E125. https://doi.org/10.1002/joc.6668

The effects of EC downscaling on SMB and other SEB variables such as melt heat energy have been investigated by Sellevold et al. (2019). Comparing four CESM1 simulations applying four different temperature lapse rates (1 K km^{-1}, 4 K km^{-1}, 6 K km^{-1}, and 9.8 K km^{-1}), the authors found that the lapse rates of 6 K km^{-}1 and 9.8 K km^{-}1 result in the most realistic SMB and melt gradients, respectively. The simulations using lapse rates of 1 K km^{-}1 and 4 K km^{-}1 performed worse. Although the study of Sellevold et al. (2019) demonstrate that a lapse rate of 6 K km^{-}1 results in a realistic SMB, the study focusses on Greenland and applies a fixed lapse rate in the same way we do. As the reviewer already mentions, lapse rates can vary both spatially and temporally over mountain ranges, such as HMA. For example, in Langtang Valley, a glaciated valley located in Nepal (SE-HMA), temperature lapse rates (based on observations) can vary between 5.2 K km^{-1} in summer and 7.6 K km^{-1} in spring (Wijngaard et al., 2019). Further it is found that temporally variable lapse rates can improve the simulation of melt and runoff in glaciated valleys (Immerzeel et al., 2014). We are therefore convinced that spatially and temporally variables lapse rates can give better results. However, the limitation is that the estimation of temporally and variable lapse rates often relies on meteorological observations, which are scarce in HMA and other mountain ranges, particularly at higher altitude.

Figure 11 – do the initial condition spin ups of the VR-CESM experiments have anything to do with why SMB starts, even in the first year, with such large losses compared to both observations/WRF? This is likely a naïve comment (not a glacier expert), but would there be any way to initialize the VR-CESM simulations to start with glacier/snow thickness comparable to observations/WRF instead of relying on the CESM spin up procedure? Or could the authors use ERA5 to spin up a standalone CLM simulation that could then provide initial conditions to the AMIP VR-CESM simulations?

For the HMA_VR7a simulation, the initial conditions spin-up could have played a role in starting with large losses compared to both observations/WRF-based outputs. However, for the HMA_VR7b simulation, we initialized the model with a snow depth of 2.5 m w.e. in every glacierized land unit. This could be similar to forcing CESM with snow thickness that is comparable with observations, except that we only applied the initialization over glacierized land units. Our experience was that in the first 5 years of the atmospheric spin-up, most of the snow over the glacierized land units already disappeared. Based on these findings, we do not think that the initial condition spin-ups have a negative impact on the SMB bias.

As for spinning up a CLM standalone simulation with ERA5, we have conducted CLM standalone simulations driven by GSWP meteorological forcings (for details we refer to our response to the second comment of Reviewer 1). As it turned out, the CLM standalone simulations forced by GSWP resulted in a more negative SMB than the HMA VR simulations, primarily due to the warmer surface climate and smaller precipitation volumes in GSWP.

Line 455 – delete "than"

"Than" is embedded in a sentence where a comparison is made between the SMB simulated over higher-altitude glaciers and the SMB simulated over "lower-altitude" glaciers. Grammatically it will therefore be incorrect to delete "than".

Line 486-501 – agreed… major bummer on the 10x SMB issue in HMA (expected SMB is -20 GT/yr, VR-CESM is, at best, -200 GT/yr). I hope the nudging exercise can help alleviate some of the large-scale biases in temperature. I hope some of my random ideas above help too.

We are very grateful for all the helpful suggestions. We also hope that the nudging experiments and experiments with improved land-atmosphere coupling schemes will improve the SMB bias and can alleviate some of the biases in temperature.

Line 546 – I'd argue most regions simulated by VR-CESM have cold biases (and across most seasons too)

We agree. We will rephrase the text to put more emphasis on the cold temperature biases.

References

Flatau, P. J., Walko, R. L., and Cotton, W. R.: Polynomial Fits to Saturation Vapor Pressure, Journal of Applied Meteorology, 31, https://doi.org/10.1175/1520-0450(1992)031<1507:pftsvp>2.0.co;2, 1992.

Gu, H., Wang, G., Yu, Z., and Mei, R.: Assessing future climate changes and extreme indicators in east and south Asia using the RegCM4 regional climate model, Clim Change, 114, https://doi.org/10.1007/s10584-012-0411-y, 2012.

He, C., Li, Q. B., Liou, K. N., Zhang, J., Qi, L., Mao, Y., Gao, M., Lu, Z., Streets, D. G., Zhang, Q., Sarin, M. M., and Ram, K.: A global 3-D CTM evaluation of black carbon in the Tibetan Plateau, Atmos Chem Phys, 14, https://doi.org/10.5194/acp-14-7091-2014, 2014.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research Atmospheres, 113, https://doi.org/10.1029/2008JD009944, 2008.

Immerzeel, W. W., Petersen, L., Ragettli, S., and Pellicciotti, F.: The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas, Water Resour Res, 50, 2212–2226, https://doi.org/10.1002/2013WR014506, 2014.

Immerzeel, W. W., Wanders, N., Lutz, A. F., Shea, J. M., and Bierkens, M. F. P.: Reconciling high altitude precipitation with glacier mass balances and runoff, Hydrol Earth Syst Sci, 12, 4755–4784, https://doi.org/10.5194/hessd-12-4755-2015, 2015.

Kang, S., Zhang, Y., Qian, Y., and Wang, H.: A review of black carbon in snow and ice and its impact on the cryosphere, https://doi.org/10.1016/j.earscirev.2020.103346, 2020.

Lalande, M., Ménégoz, M., Krinner, G., Naegeli, K., and Wunderle, S.: Climate change in the High Mountain Asia in CMIP6, Earth System Dynamics, 12, https://doi.org/10.5194/esd-12-1061-2021, 2021.

Li, D., Lu, X., Walling, D. E., Zhang, T., Steiner, J. F., Wasson, R. J., Harrison, S., Nepal, S., Nie, Y., Immerzeel, W. W., Shugar, D. H., Koppes, M., Lane, S., Zeng, Z., Sun, X., Yegorov, A., and Bolch, T.: High Mountain Asia hydropower systems threatened by climate-driven landscape instability, Nat Geosci, 15, 520–530, https://doi.org/10.1038/s41561-022-00953-y, 2022.

Liu, Y., Fang, Y., and Margulis, S. A.: Spatiotemporal distribution of seasonal snow water equivalent in High Mountain Asia from an 18-year Landsat-MODIS era snow reanalysis dataset, Cryosphere, 15, https://doi.org/10.5194/tc-15-5261-2021, 2021.

Palazzi, E., Von Hardenberg, J., Terzago, S., and Provenzale, A.: Precipitation in the Karakoram-Himalaya: a CMIP5 view, Clim Dyn, 45, 21–45, https://doi.org/10.1007/s00382-014-2341-z, 2015.

Rahimi, S., Liu, X., Wu, C., Lau, W. K., Brown, H., Wu, M., and Qian, Y.: Quantifying snow darkening and atmospheric radiative effects of black carbon and dust on the South Asian monsoon and hydrological cycle: experiments using variable-resolution CESM, Atmos Chem Phys, 19, 12025–12049, https://doi.org/10.5194/acp-19-12025-2019, 2019.

Rhoades, A. M., Ullrich, P. A., Zarzycki, C. M., Johansen, H., Margulis, S. A., Morrison, H., Xu, Z., and Collins, W. D.: Sensitivity of Mountain Hydroclimate Simulations in Variable-Resolution CESM to Microphysics and Horizontal Resolution, J Adv Model Earth Syst, 10, 1357–1380, https://doi.org/10.1029/2018MS001326, 2018.

Shea, J. M., Immerzeel, W. W., Wagnon, P., Vincent, C., and Bajracharya, S. R.: Modelling glacier change in the Everest region, Nepal Himalaya, Cryosphere, 9, 1105–1128, https://doi.org/10.5194/tc-9-1105-2015, 2015.

Slater, A. G., Schlosser, C. A., Desborough, C. E., Pitman, A. J., Henderson-Sellers, A., Robock, A., Vinnikov, K. Y., Mitchell, K., Boone, A., Braden, H., Chen, F., Cox, P. M., De Rosnay, P., Dickinson, R. E., Dai, Y. J., Duan, Q., Entin, J., Etchevers, P., Gedney, N., Gusev, Y. M., Habets, F., Kim, J., Koren, V., Kowalczyk, E. A., Nasonova, O. N., Noilhan, J., Schaake, S., Shmakin, A. B., Smirnova, T. G., Verseghy, D., Wetzel, P., Xue, Y., Yang, Z. L., and Zeng, Q.: The representation of snow in land surface schemes: Results from PILPS 2(d), J Hydrometeorol, 2, https://doi.org/10.1175/1525-7541(2001)002<0007:TROSIL>2.0.CO;2, 2001.

Wang, X., Tolksdorf, V., Otto, M., and Scherer, D.: WRF-based dynamical downscaling of ERA5 reanalysis data for High Mountain Asia: Towards a new version of the High Asia Refined analysis, International Journal of Climatology, 41, 743–762, https://doi.org/10.1002/joc.6686, 2021.

Wijngaard, R. R., Steiner, J. F., Kraaijenbrink, P. D. A., Klug, C., Adhikari, S., Banerjee, A., Pellicciotti, F., van Beek, L. P. H., Bierkens, M. F. P., Lutz, A. F., and Immerzeel, W. W.: Modeling the response of the langtang glacier and the hintereisferner to a changing climate since the little ice age, Front Earth Sci (Lausanne), 7, https://doi.org/10.3389/feart.2019.00143, 2019.