We thank the editor for providing more comments to improve our manuscript. Below are our detailed replies to these comments.

Comment: Thank you for revising the manuscript and for your responses. Since one point of critique in the first round (open discussion phase) was serious ("insufficient model validation") and the revised MS has changed quite substantially, I decided to ask a third, new referee for an independent opinion. This person has a strong background in regional atmospheric modeling. You will find his/her review attached, and you will see that the referee has major concerns with the study and rated the scientific quality of the paper as "poor".

This referee points out a "lack of realism", which links directly to the aforementioned "insufficient model validation", suggesting that the paper has not improved convincingly in that respect. In particular, the uncertainty in precipitation is emphasized by the new referee, which also connects to Dieter Scherer's comment "...the entire study depends on the accuracy of downscaled precipitation. It would therefore be of utmost interest to better understand the uncertainties in the WRF output." from the first round of review. In addition, my own review (see attached) raises a question in the same direction (are model and observations in agreement? See point **(2)** of my comments). Therefore, revisions are necessary, and I will make a final decision after the re-submission and your responses whether the manuscript can be considered for publication or not.

Reply: We agree that there are unquantified uncertainties in our model, which we already acknowledge in our manuscript. However, many uncertainties are simply unquantifiable due to the lack of data in WKSK, as we also indicate in our manuscript. This is a problem that is relevant for all studies in this region, including those already published. We disagree that our modelling lacks realism, since all our models are ultimately driven by reanalysis data, which are all largely influenced by observations of the surface and atmosphere. However, we now also added further validation with remote sensing data, showing that the atmospheric moisture in our model excellently reproduces observations, much in contrast to the statements of the referee. Using a combination of techniques, our study is the first to reasonably reproduce the pattern of mass balances in HMA and we provide many new insights into what might cause this anomalous mass balances. Furthermore, our paper gives useful new avenues for future research. In summary, we think our paper is very much suitable for publication in The Cryosphere. Although our title is a concise summary of our results, we now tried to better acknowledge the uncertainties in our modelling, by changing the title as follows: **Towards understanding the pattern of glacier mass balances in High Mountain Asia using regional climatic modelling**. We also add to the abstract: "…we reproduce the observed patterns of glacier mass balance in High Mountain Asia of the last decades **within uncertainties**." The rest of the abstract already has cautious language.

To gain confidence in our modelling results we have performed comparisons and validation efforts that are well beyond what is normally done in similar studies in the published literature. We have:

1) Compared our model results with station data, which is what is commonly done, but is challenging due to the scale difference and large uncertainties in precipitation observations in mountains.
2) Compared our model results with several reanalysis datasets, which is commonly done in separate papers.
3) Compared different ET datasets, which is normally not done and is normally extended to a study in itself.
4) Compared our model results with AIRS+AMSU retrievals, which provides a new kind of validation that is not commonly done.

All of our comparisons show that our model compares at least reasonably well, and often very well, with observations for all of the most relevant parameters, and hence certainly does not "lack realism". We also already acknowledge and discuss our model uncertainties to a larger extent than most atmospheric modelling papers in the region, and even globally.

However, the statement about the lack of realism possibly does not only relate to our model validation, but with a misinterpretation of our methods by this referee. We understand there would certainly be a lack of realism if we would have used the WRF fields directly to model the mass balance, or the mass balance gradient, which is what the referee might have assumed. We now provide a detailed discussion of our approach of using relative changes in temperature and precipitation and now demonstrate that our method is much more robust against the potentially large biases in the WRF model (e.g. by the reduced resolution or choice of physics modules), and is hence much more realistic than downscaling the WRF data directly. We also further discuss the uncertainties of our method. We add:

… We show that such a pattern can be reproduced using relative changes in temperature and precipitation in recent decades. Since we used relative changes to force our glacier model, we are less influenced by errors in the absolute precipitation amounts, caused by our low resolution or by our choice of model physics. We illustrate this using WRF runs performed for de Kok et al. (2018) for May-September of two years. We ran WRF at two resolutions: at 20 km with same the physics settings as in this study, but without any nudging, and at 4 km, which is of high enough resolution to explicitly resolve convection and avoid the cumulus parameterisation. There are large local differences in precipitation between the two runs, mainly due to the difference in resolution. However, when the relative ratio of the precipitation is plotted for two years (Fig. 16), similar to what is used in the glacier model, the two set-ups give much more similar patterns. Snowfall gives very similar results, but we decided to show total precipitation, where total numbers and cumulus errors are expected to be even higher. The relative changes in precipitation do not markedly show the topography, in contrast to the individual precipitation fields. Rather, relatively large regions show similar interannual changes in the precipitation. The patterns of precipitation change also agree well between the 20 km results and the 4 km results, despite the very different treatment of the convection and the difference in topographic resolution. The differences between the scaling factor in the two cases can be of the order of tens of percent, which is much smaller than the difference in absolute precipitation amounts that would be needed to model the mass balance directly from the WRF fields. Also temperatures are mutually correlated over larger areas in WKSK (e.g. Forsythe et al., 2017) and the glacier mass balances in HMA also vary mainly over a large scale, suggesting that large-scale weather patterns are on average more important in controlling the interannual variability of temperature and precipitation than the differences between valleys. The use of relative changes in temperature and precipitation has thus made our results more robust against possible errors in the detailed treatment of the complex mountain meteorology.

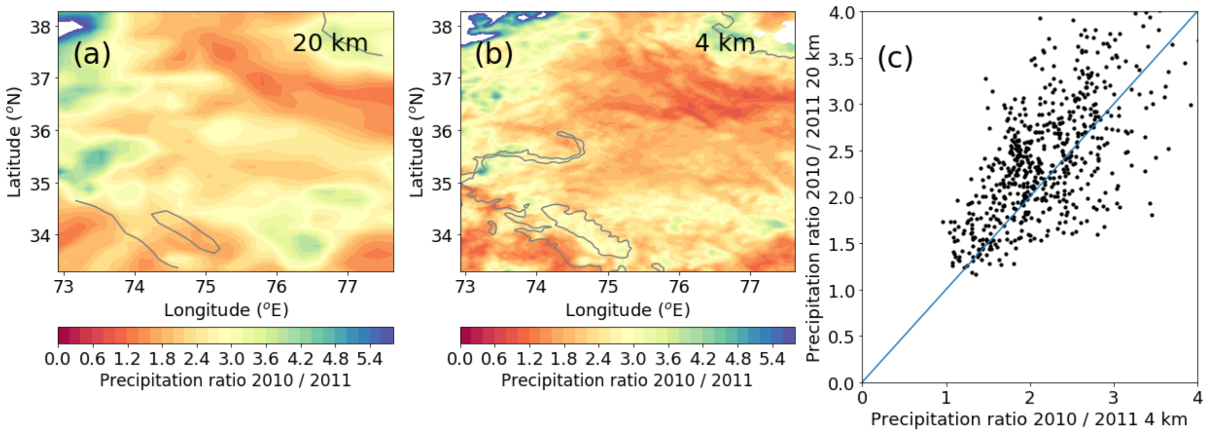


Figure 16: Precipitation ratios between May-September of two years for the WRF run at 20 km, with cumulus parameterisation (a), that at 4 km resolution, without parameterisation (b), and the two compared, when binned at the resolution of the 20 km run (c). The 2000 m elevation contour is indicated by a solid line.

One of our main sources of error is setting up the initial mass balance gradient, and our assumption that the glaciers are initially in balance. Due to the inertia of the glaciers, the initial condition has relatively large influence on the eventual mass balance decades later, as discussed above. Furthermore, any errors in the mass balance gradient, e.g. due to errors in the downscaling of ERA-Interim data, will affect the temperature and precipitation sensitivities presented here, but will have less impact on the overall pattern of mass balances in HMA, since they are mostly determined by the changes of temperature and precipitation.

Comment: Below I list ~20 examples that I caught during my reading, which illustrate a lack of precision in statements or procedures. While each case on its own is probably a minor problem, all together make reading the paper quite hard. Please go through the paper carefully and revise these and similar problems to enhance the unambiguity of statements. I hope you can see that the below examples are not reader-friendly.

Reply: Although the editor and all three referees have rated the presentation of the results as "good" or "excellent" in previous rounds, we agree that there is always room for further improvement and we have taken all the points into consideration.

Comment: 96-100: Several instances of "temperature"; it is unclear whether they refer to air or the glacier (surface).

Reply: We now included "air" before "temperature" six times in this paragraph.

Comment: 101-105: "snow" and "snowfall" are used here; do they refer to the same (I guess you mean snowfall as you refer to solid precipitation)?

Reply: We replaced all instance there to say "snowfall".

Comment: 123: "amount of water"; all phases or just a particular one?

Reply: This is indeed all phases of water. We add: "…the **total** amount of **all** water in the parcel…"

Comment: 181: RMSE of 1.8 °C; were annual, seasonal, monthly values used in the calculation? It is not obvious.

Reply: We add: "…the stations **for the time-series of seasonal mean temperatures** is 1.8°C."

Comment: Figure 2 caption: What is the "nearest WRF grid" in this context? Do you mean "nearest grid point"?

Reply: We now added "point".

Comment: Figure 2: The figure is supposed to show a station/WRF comparison. How do panels (b) and (e) fit here? They show a station trend.

Reply: The trends give an idea about the content of the station data. We elaborate on this in the text: "The stations generally indicate a strong heating trend, **but also show relatively large differences for close-by stations**." and "**The stations show an increasing trend in May-September precipitation in the western Tarim basin and most of the eastern Tibetan Plateau.**"

Comment: 198: RMSE of 11.4 mm/month; were annual, seasonal, monthly values used in the calculation? Again, it is not obvious (per month can also be used as unit for mean seasonal or annual values).

Reply: We add: "…**for the time-series of seasonal mean precipitation** is 11.4 mm per month."

Comment: Figure 2: it is referred to a lot in the text, but almost always without specification of which panel is meant (a few times I understand that the entire figure is addressed, but that can't be the case always).

Reply: In the text, where the figure is first presented, different panels are discussed, but the panels show very different variables, which we clearly mention in the text. We now also add a reference to the figure panel each time. Besides the text that directly presents Fig. 2, we only find one direct other reference to it in the text. To further clarify, we add the panel number when discussing Fig. 3: "The magnitudes of the trends are also generally smaller than those in the station data (Fig. 2**b**)."

Comment: 232: is "figure 3 in de Kok et al., 2018" meant? Figure 3 in the present paper is a temperature figure.

Reply: This is indeed the case. Although the original text is a common format for citing page numbers or figures in other papers, we rephrase now as: "(**Fig. 3 of** de Kok et al., 2018)."

Comment: Figure 3: Time step for the correlation calculation not clearly specified (annual, season, …).

Reply: All correlation calculations are performed in a similar way, and we now elaborate in Section 2.4: "Pearson correlation coefficients are calculated **between pairs of different datasets (e.g Figs. 2-5)** using the vectors of annual or seasonal mean values, **with one value for each year**. **The figures indicate over which period the mean is taken for each year.** The trends shown in Figs. 2-5, 8, and 17 are the slopes from linear fits to **these** vectors.

Comment: Figure 4: Time step for the correlation calculation not clearly specified (annual, season, …).

Reply: All correlation calculations are performed in a similar way, and we now elaborate in Section 2.4: "Pearson correlation coefficients are calculated **between pairs of different datasets (e.g. Figs. 2-5)** using the vectors of annual or seasonal mean values, **with one value for each year**. **The figures indicate over which period the mean is taken for each year.** The trends shown in Figs. 2-5, 8, and 17 are the slopes from linear fits to **these** vectors. "

Comment: 279-281: How can changes (increase/decrease) be inferred from Figure 8, which shows the mean diurnal cycles?

Reply: We are confused by the comment of the editor here. Nowhere in the paper do we show the diurnal cycle. We do show the mean seasonal cycle in Fig. 8, but we also show the trends in panel d. We now mention this panel explicitly.

Comment: Figure 7: The captions says regions with "growing" and "shrinking" glaciers, but the legend shows WKSK and SW HMA. Are these two definitions exactly the same regions?

Reply: The two areas are sub-areas of WKSK and SW HMA, which we now clarify in the caption: "…for two nearby 2x3° bins **in WKSK and southwestern HMA** that have…"

Comment: Figure 9: Is the scale bar adjusted to the min/max values in the maps? I can't see much dark blue in the maps.

Reply: The positive part of the scale bar follows those of the much-read works of Kääb et al. and Brun et al., which allows for better comparison with these works.

Comment: 332: "snow" or "snowfall"; I guess you mean the latter?

Reply: We now write "snowfall"

Comment: 382-385: Same as above, which time step is used for calculating the correlations? "Interannual" can also compare seasons or months between years (or annual values).

Reply: All correlation calculations are performed in a similar way, and we now elaborate in Section 2.4: "Pearson correlation coefficients are calculated **between pairs of different datasets (e.g. Figs. 2-5)** using the vectors of annual or seasonal mean values, **with one value for each year**. **The figures indicate over which period the mean is taken for each year.** The trends shown in Figs. 2-5, 8, and 17 are the slopes from linear fits to **these** vectors. "

Comment: Discussion starting in 409: "roughly matches" is not clear enough. Also, your Figure 6b shows snowfall trends, yet you refer to precipitation here. The cited reference also shows precipitation. Please be consistent for comparisons.

Reply: We now attempted to clarify the section, including adding an extra panel, showing the WRF JAS total precipitation trend. We now write:

The pattern of **snowfall** trends in Fig. 8b roughly matches the **precipitation** pattern that is expected from an increasing influence of summer westerlies, as shown by Mölg et al. (2017). **From this similarity, one could wonder whether the snowfall pattern from Fig. 6b is mainly caused by summer westerlies.** These **summer** westerlies are also associated with strong heating and drying trends of the Indus Basin. An increase in irrigation also produces a very similar precipitation pattern

as the pattern for summer westerlies, yet causes a cooling and wetting of the Indus Basin (de Kok et al., 2018). Our JAS trends of near-surface temperature and specific humidity **from WRF** (Fig. 17) indicate mostly cooling and wetting trends **in the Indus basin**, which is more in line with the increase in irrigation than with the increase in summer westerlies. ERA5 data for JJA also indicates a similarly strong irrigation effect in the Indus basin (Farinotti et al., 2020), **as indicated by a wetting and cooling trend**. The moisture tracking results (Figs. 14 and 15) indicate that much of the additional snowfall occurs in spring and summer, and originates from the East, with a large role for the irrigated areas. The decrease in precipitation in southwestern HMA is also clearly associated with westerly winds in winter, but not those in summer (see Figs. 10d and 14c). The pattern of **snowfall** trends in Fig. 8b is **thus** not only the result of changes in summer. **When only JAS is considered, the pattern of precipitation trends look different from the annual snowfall trends (Fig. 17b). Therefore, the summer westerlies are likely not the main driver for the snowfall pattern seen in Fig. 8b.** However, the May westerlies clearly have an important role in transporting the increase in evaporation from the Caspian Sea (Chen et al., 2017) to WKSK. Besides the Caspian Sea, the westerlies are mainly associated with a decrease in snowfall when the whole year is considered (Fig. 14a).

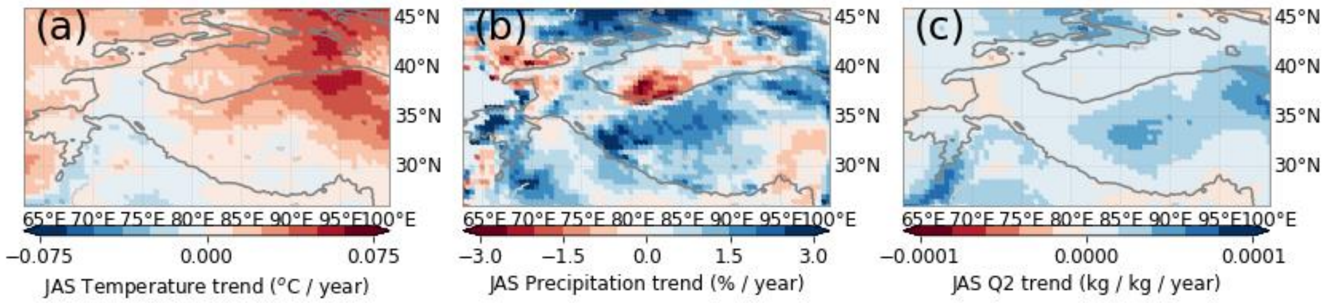


Figure 17: WRF trends between 1980-2010 of near-surface temperature (a), total precipitation (b), and specific humidity (b) between July-September, averaged over 0.5x0.5° bins for clarity. The 2000 m elevation contour is indicated by a solid line.

Comment: General: What tests are used to determine the p values?

Reply: We add: **"P-values for the correlations are determined using the beta function, as implemented in SciPy (Virtanen et al., 2020)"**

Comment: General: Are the trends tested for significance?

Reply: We tested statistical significance before, but we found that it is actually not at all relevant for our conclusions, since in many cases the variability is dominated by real-life variability, not by measurement error. So, for instance, suppose we know the real snowfall to infinite precision. Even then, any trend might not be significant if there is large (interannual) variability. Yet, the trend would be very much real in this case, since we know the snowfall with infinite precision. Hence, the statistical significance is mainly an indirect measure of the size of the trend compared to the variability, which is not relevant for our study. We think that stating the significance (several positive snowfall trends in Kunlun Shan are indeed statistically significant) would be misleading. The thing that one would like to know is the robustness of the trends under the given model uncertainty, but this would require separate studies, such as the ensemble approach we propose in the discussion.

Comment: General: with the addition of new data sets, mixing up Methods and Results has become more serious than in the first version. The readability would benefit from having more descriptions of data and technical procedures before the results section (Section 3).

Reply: There is indeed something to be said for this, but we think it is also good to separate the data that is produced for this paper from the external datasets. Effectively, almost all the “new” datasets are all included in Section 3.1, before the presentation of our own results, and hence the mixing is mainly limited to this one subsection.

Comment: **(2) Scientific Contents**

My main topical comment refers to Section 3.3, where one key message is that areas of growing or shrinking glaciers are consistent in model and observations. While this would be a nice result, I assume that readers will have trouble understanding it when they look at Figure 10. In particular, where the model region tends to be positive (marked Box 1 below) the Brun values are mostly negative, and where the model region tends to be negative (Box 2) the Brun results tend to be positive. One could also conclude that model and observation show the opposite with regard to neutral/stable mass balances. This discrepancy adds to the referee assessment of “a lack of realism”.

Reply: We agree that, if one would only look at the figure as it is, without reading the text, one might indeed come to a different conclusion. However, our assessment of the figure takes into consideration the large error bars on the measurements, which we do mention in the text. We now include an illustration of the size of the error bar to show better that there is indeed a large consistency. Naturally, the large size of the error bars means that there are little constraints to validate the model outcome, but this is simply a reflection of our lack of knowledge regarding the Karakoram anomaly.

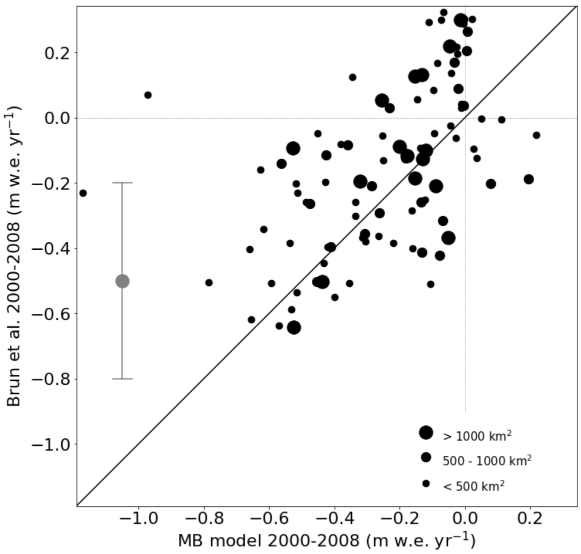


Figure 12: Comparison between mean modelled mass balances from this work, binned on a 1x1° grid as in Fig. 11, and those derived from observations of Brun et al. (2017), which are on the same grid, between 2000-2008. The size of the mean errors on the observed mass balances is illustrated by the grey error bar.

We thank the new referee for a very detailed and critical look at our manuscript. Below are our detailed replies to the comments made by the referee.

Comment: Employing a simple glacier mass balance model forced with dynamically refined ERA-Interim at 20km with irrigation, the study mainly focuses on reproducing the observed mass balance gradients of the WKSK glaciers, as the role of irrigation on summer snowfall increase as well as the impact of increasing westerly solid precipitation both resulting in positive or balanced mass balances have already been reported either by the authors or other studies. I very much like the idea of the study.

Quite common among the most of climate modelling studies focusing the WKSK region is the lack of their validation against (high-altitude representative) observations within the WKSK region itself, and especially against those that actually include snow, although such observations are quite a few. Lacking such validation initially, revised study now includes more stations, adjust WRF-20km temperatures prior to their comparison with the stations, and also, validates the WRF-20km against the ERA5, ERA5-Land and HAR fine-scaled reanalysis datasets.

Besides the above cited additions and revisions in the study, I am of the view that although observed mass balances of the WKSK glaciers are reproduced - at least to some degree of satisfaction - however the path to such outcome seriously lacks realism. There are large and unquantified uncertainties at each step of reproducing the delicate mass balance gradient of the WKSK glaciers, which need to be first reduced to the extent possible, quantified and then assessed for their possible impacts on the results and conclusions. My particular concerns are given below:

Reply: We agree that there are unquantified uncertainties in our model, which we already acknowledge in our manuscript. However, many uncertainties are simply unquantifiable due to the lack of data in WKSK, as we already indicate. We disagree that the modelling lacks realism, since all our models are ultimately driven by reanalysis data, which are all largely influenced by observations of the surface and atmosphere. We have now added more validation and now demonstrate that our approach of using relative changes in temperature and precipitation greatly reduces potential errors in understanding the delicate mass balances. We give details of our arguments and changes below.

MAJOR CONCERNS:

Comment: To establish the robustness of the WRF-20km simulations for further use in mass balance modelling of the WKSK glaciers, its scale should be refined and its validation must be performed over the complex terrain of the focused regions against few available high-altitude observations from the WKSK region, instead of only against the stations from non-glaciated surrounding areas. Because relative to the less complex terrain of the surrounding areas, coarse grid size dynamical refinement (as in the study) perform poorly over highly concentrated topography of the western Tibetan Plateau, Karakoram and adjacent regions, irrespective of the forcing datasets, featuring substantial cold (6-10deg) and wet biases. These biases are spatially heterogeneous and are difficult to adjust statistically, particularly for precipitation. Not surprising that in the revised study, the WRF-20km temperatures have been adjusted only for comparison against the station observations but not for calculating melt season temperatures for mass balance modelling, whereas, precipitation was not adjusted at all. Substantial cold biases at 20km grid size over complex terrain may result in overall shorter melt season, reduced energy available for melt and anomalous snowfall amounts. Reducing these biases and then quantifying the effect of remaining biases on the results are therefore fundamental to establish the robustness of climatic simulation, and in turn, of modelling delicate mass balance gradients of the WKSK glaciers. Establishing the robustness of climatic simulation requires at least introducing a convection-permitting scale and resolving valley scale physical processes explicitly, by introducing a further model nest spanning over the WKSK region and then extensively validating it against available high-altitude stations from SIHP WAPDA, PMD, Agha Khan Agency for Habitat, CMA, EVK2-CNR, and others from the WKSK region at least within the 2000-2010 period, which has been used for mass balance comparison.

Further, precipitation in the only used model realization is highly sensitive to the chosen cumulus parameterization, microphysics and related physics in the climate model. Any change in physics leads to significantly different magnitudes of precipitation and signs of change, besides other climatic conditions. An effort to quantify the sensitivity of results to chosen physics in the model is missing. Recommendation from the literature could have been useful too. Chosen model physics is actually based on de Kok et al., 2018, who state that their 7-year simulation was not aimed at accurately reproducing the reality but only to show the effect of irrigation, unlike this study.

Reply: We agree that one would need to downscale and bias-correct the WRF data if we would have used the data directly to model the glacier mass balance, or determine the spatially resolved mass balance gradient from such input. We also agree that the choice of parameterization can greatly change the modelled precipitation amounts. However, such biases in precipitation or temperature are not as relevant as the referee suggests, since we do not calculate glacier mass balances directly from the WRF fields. As stated in the methods, our glacier mass balance outcomes are the result of using *relative* temperature and precipitation changes. This is in contrast to most other work, such as Collier et al. (TC, 2013) and Kumar et al. (GRL, 2015). Our initial mass balance gradient is calculated by assuming a glacier in balance, which is an assumption we already discuss in our manuscript. Initially, ERA-Interim data is used to determine the mass balance gradient, where temperature is downscaled/"corrected" to calculate degree days, as already stated in the methods. Any error in the determined mass balance gradient will mainly result in a change in climate sensitivity, and hence will change the values of the mass balances somewhat, but will not change the pattern of positive-negative mass balances much, which comes from the relative changes of temperature and precipitation from WRF. To illustrate and discuss these points we now add more text in the discussion, which includes comparison of convection-permitting WRF runs at 4 km resolution with WRF runs at 20 km resolution:

… We show that such a pattern can be reproduced using relative changes in temperature and precipitation in recent decades. Since we used relative changes to force our glacier model, we are less influenced by errors in the absolute precipitation amounts, caused by our low resolution or by our choice of model physics. We illustrate this using WRF runs performed for de Kok et al. (2018) for May-September of two years. We ran WRF at two resolutions: at 20 km with same the physics settings as in this study, but without any nudging, and at 4 km, which is of high enough resolution to explicitly resolve convection and avoid the cumulus parameterisation. There are large local differences in precipitation between the two runs, mainly due to the difference in resolution. However, when the relative ratio of the precipitation is plotted for two years (Fig. 16), similar to what is used in the glacier model, the two set-ups give much more similar patterns. Snowfall gives very similar results, but we decided to show total precipitation, where total numbers and cumulus errors are expected to be even higher. The relative changes in precipitation do not markedly show the topography, in contrast to the individual precipitation fields. Rather, relatively large regions show similar interannual changes in the precipitation. The patterns of precipitation change also

agree well between the 20 km results and the 4 km results, despite the very different treatment of the convection and the difference in topographic resolution. The differences between the scaling factor in the two cases can be of the order of tens of percent, which is much smaller than the difference in absolute precipitation amounts that would be needed to model the mass balance directly from the WRF fields. Also temperatures are mutually correlated over larger areas in WKSK (e.g. Forsythe et al., 2017) and the glacier mass balances in HMA also vary mainly over a large scale, suggesting that large-scale weather patterns are on average more important in controlling the interannual variability of temperature and precipitation than the differences between valleys. The use of relative changes in temperature and precipitation has thus made our results more robust against possible errors in the detailed treatment of the complex mountain meteorology.

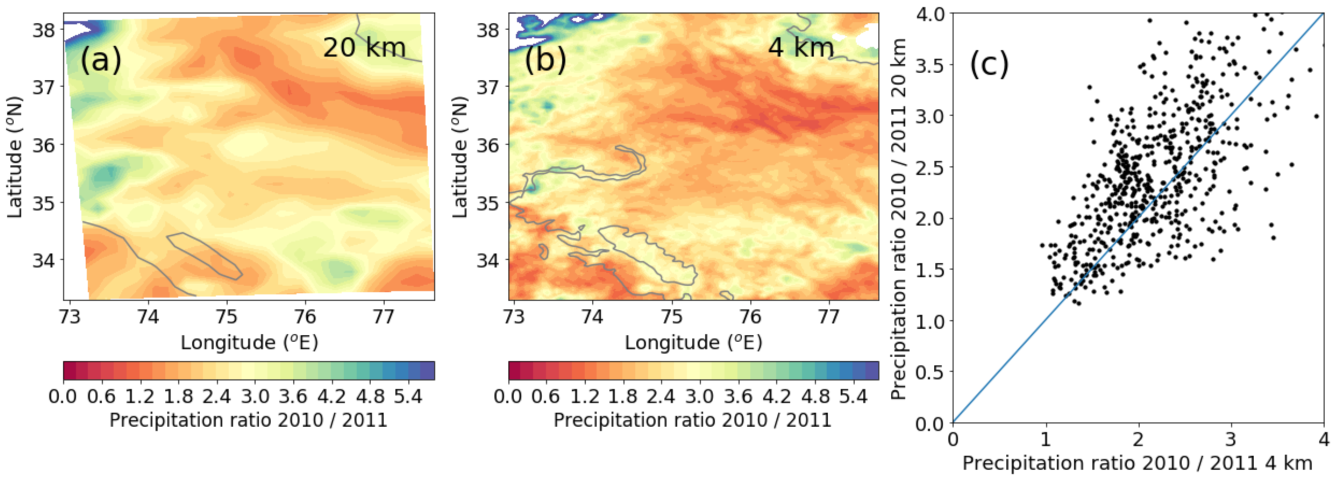


Figure 16: Precipitation ratios between May-September of two years for the WRF run at 20 km, with cumulus parameterisation (a), that at 4 km resolution, without parameterisation (b), and the two compared, when binned at the resolution of the 20 km run (c). The 2000 m elevation contour is indicated by a solid line.

One of our main sources of error is setting up the initial mass balance gradient, and our assumption that the glaciers are initially in balance. Due to the inertia of the glaciers, the initial condition has relatively large influence on the eventual mass balance decades later, as discussed above. Furthermore, any errors in the mass balance gradient, e.g. due to errors in the downscaling of ERA-Interim data, will affect the temperature and precipitation sensitivities presented here, but will have less impact on the overall pattern of mass balances in HMA, since they are mostly determined by the changes of temperature and precipitation.

The further nesting of our WRF results could indeed be useful in future studies, if one wants to determine mass balances directly from the WRF fields, e.g. as we did in Bonekamp et al. (Frontiers, 2019). However, this would be a different study and would require work that is clearly beyond a single paper such as our manuscript, since downscaling such a large region to a resolution of 1 km for 30 years is computationally too expensive at the moment. Furthermore, the data that the referee mentions could be useful for further validation of regions outside the region where the Karakoram anomaly is present, where the data is taken. However, since most of these data are not easily accessible, and our focus is mainly on the anomalous glaciers in WKSK, we do not see how this will make the paper much stronger beyond the work that is already performed. A good agreement in e.g.

the Himalaya is also no guarantee of a good results in WKSK. For instance, Kumar et al. (Scientific Reports, 2019) show excellent agreement at the validation sites, but show very positive mass balances in the Karakoram, and negative mass balances in Kunlun Shan, which is in contrast to what is observed.

Our model physics was indeed the same as in de Kok et al. (2018), but is actually based on the work of Collier and Immerzeel (2015) and Bonekamp et al. (2018), who do aim to reproduce reality. We now add: "…**which are based on the work of Collier and Immerzeel (2015) and Bonekamp et al. (2018)…"**

Comment: The implementation of irrigation in 20 km WRF simulation lacks realism as it perturbs precipitation at each timestep based on monthly crop water demand rate calculated by PCR_GLOBWB forced by different dataset (may be CRU TS2.1 and ERA-40), and most importantly, ignoring the on-ground facts, such as deficit conditions and the irrigation efficiency. This can lead to anomalously higher moisture availability that yields increased snowfall in the neighboring regions and subsequently can positively affects the mass balance results. Hence, it is important to realistically implement the irrigation in the model to avoid introducing spurious atmospheric moisture amounts that are favorable to the conclusions of the study. Qualitative validation against non-validated proxy evapotranspiration observations does not add to the robustness of the WRF-20km irrigation and requires to be replaced with quantitative validation against reliable datasets including quantification of their uncertainty and subsequent impacts on the conclusions.

Reply: We agree that it important that it is important not to overestimate the irrigation, which is why we chose to supply a fixed amount of water, instead of e.g. keeping soil moisture constant, which can potentially greatly overestimate the irrigation in locations such as the Tarim basin. Unfortunately, real irrigation gifts are largely unknown. PCR-GLOBWB is forced by WATCH data, which is based on ERA-Interim, with a bias-correction based on CRU. We add to the methods: "**The PCR-GLOBWB runs were forced by WATCH data, based on ERA-Interim (Weedon et al., 2014)"** Because of the uncertainty in the irrigation, it would be useful to compare different forcings and irrigation schemes, but this is well beyond the scope of the current paper. However, we now do verify the atmospheric moisture amounts at the irrigated area in the Tarim basin, which is the most crucial region in our study, using AIRS and AMSU data. These are retrievals that directly infer the humidity from infrared and microwave radiances at different wavelengths, with practically no major model assumptions. Since humidity is an important driver for precipitation, which is harder to measure, the AIRS and AMSU retrievals provide a good validation dataset for our WRF model. The results show that our model shows very close agreement with remotely determined water vapour, which contrasts the proposed lack of realism. We add:

We further investigate the realism of the effect of irrigation in our model by comparing remotely sensed surface specific humidity from AIRS and AMSU retrieval with our WRF specific humidity at 2 metres. These are not exactly the same quantities, as AIRS has a finite vertical resolution, but the variations over time can be compared. We focus on the irrigated area in the Tarim basin, close to the Kunlun Shan, which is the most important in the later discussion on the Karakoram anomaly. The flat terrain makes the retrievals near the surface more certain compared to mountainous regions, where altitude, and hence pressure and humidity, strongly vary within the spatial resolution of the measurements. The comparison between means over May-September for 2003-2010 is shown in Fig. 7. Even though we did not nudge WRF towards ERA-Interim near the surface, the model still follows the humidity observations in the irrigated region in the Tarim basin very

closely, with a Pearson correlation coefficient of 0.97. This gives further confidence that the irrigation we apply there is not unrealistic.

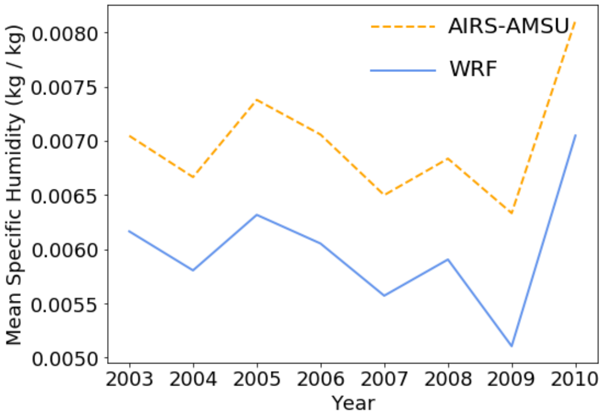


Figure 7: May-September mean specific humidity at 2 metres from WRF (blue, solid line) and AIRS-AMSU surface humidity (orange, dashed line) for a 1° x 1° bin around 38.5° N, 77.5° E, which is an irrigated area in the Tarim that contributes to the snowfall in WKSK.

Furthermore, we now also add a comparison with AIRS-AMSU retrievals for moisture fields above the mountains, which shows that our atmospheric moisture certainly does not lack realism. We add:

The station and reanalysis data show a good agreement with our WRF output in many locations, but the comparison is hampered in WKSK due to the aforementioned fundamental uncertainties. Remote sensing data can also be used for comparison, but also there, uncertainties can be very high. This is especially true for precipitation measurements in mountainous areas, but also other remote surface measurements of relevant parameters are uncertain in mountainous areas (Lundquist et al., 2019). However, the atmosphere above the mountains can be measured with some confidence. Especially the atmospheric humidity can be used to increase the confidence in the interannual variability of the precipitation, since the two are strongly related. Here, we compare retrieved atmospheric humidity from AIRS and AMSU data (AIRS Science Team and Teixeira, 2013) above the mountains with humidity from our WRF output. These retrievals determine the humidity from satellite measurements at wavelengths in the infrared and microwave with very limited assumptions (Susskind et al., 2014), and hence can be considered as good validation dataset. Figure 5 shows the comparison of the mean AIRS specific humidity between May-September and between 400-500 hPa, and the corresponding WRF specific humidity interpolated at the middle of this layer (447.2 hPa) for the overlapping years 2003-2010, binned at the AIRS-AMSU resolution. The two datasets show a very high overall agreement, both in the patterns of humidity trends, as well as the correlation of interannual variability, with correlation coefficients generally above 0.9. This analysis shows that the moisture transport in our WRF model closely follows what we know of the atmosphere around WKSK. Near the edges of our modelling domain, our errors are naturally larger.

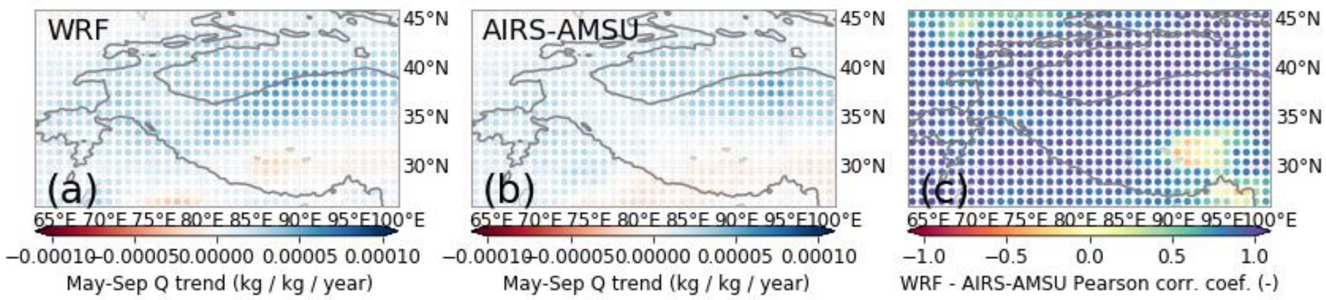


Figure 5: May-September mean specific humidity trends at 447.2 hPa for WRF (a), and between 400-500 hPa for AIRS-AMSU (b) between 2003-2010. Panel (c) shows the Pearson correlation coefficient between them.

Other comments:

• What is the statistical significance of presented trends in temperature and snowfall? Do presented slopes actually carry any physical meanings?

Reply: We tested statistical significance before, but we found that it is actually not at all relevant for our conclusions. Importantly, the statistical significance of a climatic trend has nothing at all to say about the "carrying of physical meanings", since in many cases the variability is dominated by real-life variability, not by measurement error. So, for instance, suppose we know the real snowfall to infinite precision. Even then, any trend might not be significant if there is large (interannual) variability. Yet, the trend would be very much real in this case, since we know the snowfall with infinite precision. Hence, the statistical significance is mainly an indirect measure of the size of the trend compared to the variability, which is not relevant for our study. We think that stating the significance (several positive snowfall trends in Kunlun Shan are indeed statistically significant) would be misleading. The thing that one would like to know is the robustness of the trends under the given model uncertainty, but this would require separate studies, such as the ensemble approach we propose in the discussion.

• Line 256-257: how it is established that GLEAM provides unrealistically low evapotranspiration in heavily irrigated arid regions in July. Validation of WRF-20km against the employed evapotranspiration proxy observations is completely subjective and unreliable. In fact, landing in the middle of multiple non-validated unreliable datasets does not establish the robustness of the WRF-20km simulations.

Reply: It is already clear from this simple example that GLEAM underestimates the evotranspiration in e.g. the Tarim basin, since the values of the dry desert and irrigated agriculture are almost identical. Common sense tells us that this is very unlikely in reality. We now add more explicitly: "However, it is clear that GLEAM does not represent the irrigated areas well, with evapotranspiration in heavily irrigated arid regions in July **that is as low as the surrounding deserts in e.g. Tarim and Indus basins, which is not realistic**."

This comparison was also not meant as a robust validation, and we do not claim it to be. It does illustrate the problem of lack of reliable datasets. All of these datasets are in fact validated in some way, but that does not always guarantee accuracy in all areas. Furthermore, one can hardly describe ERA-Interim as an "unreliable dataset", since it has been validated and compared with measurements countless times. We now add: "These datasets **are all validated to some extent, but** vary greatly **nevertheless**, as we illustrate in Fig. 5 for July 2010.", and: "**However, the figure illustrates the problem with the high uncertainty in evapotranspiration over large areas in and**

around HMA."

• I think, unlike west Kunlun Shan, Karakoram region feature accumulation during winter and spring. Validation of winter and spring climatic conditions seems important here.

Reply: In this paper we mainly focus on the Karakoram anomaly and hence more on the Kunlun Shan. A too strong focus on the Karakoram has maybe hampered understanding of the Karakoram anomaly in the past. The whole area of our interest shows a peak in precipitation, and its increase, between May-September (old Figure 8) and hence we mainly focus on this. Furthermore, station comparisons in winter and spring are also hampered by the fact that there is more snowfall, which is very difficult to measure accurately.

• Is the water demand calculated based on ERA-Interim? I guess that the water demand calculated by the PCR-GLOBWB was based on CRU/ERA-40 datasets, which are different than those used here. If yes, any explanation on the effects on results should be added.

Reply: The PCR-GLOBWB runs we used were forced by WATCH data, which are based on ERA-Interim, and have a bias correction based on CRU. We add to the description of the Methods: "**The PCR-GLOBWB runs were forced by WATCH data, based on ERA-Interim (Weedon et al., 2014)."**

• How precipitation at each time step was perturbed in the model is not clear. How it has been achieved?

Reply: We add the following clarification: "**The precipitation is added to the NOAH-MP surface module, but the atmospheric module is not altered."**

• Implementing irrigation through continuous light rain in the study completely ignores the significant impact of irrigation timing on the climate and seasonality and the state of vegetation. I think direct perturbation of soil moisture is a better approach that imitates irrigation via flooding of the surface and disregards other reservoirs such as the canopy layer. Hence, it is important to know what effect introducing of continuous precipitation had on realistically reproducing surface parameters? I hope gridded observations at least over the plain areas of south Asia are representative and can be used for validation.

Reply: It was shown in previous work that saturating the soil can greatly overestimate the irrigation. This would presumably be especially relevant for the extremely arid Tarim basin. Above, we show that we now added a comparison with AIRS-AMSU data in the Tarim basin, which is the most relevant irrigated region for our study. A more sophisticated irrigation model that better includes the daily cycle can indeed be used in future work, but the excellent agreement with AIRS-AMSU data indicates that the interannual variation of atmospheric moisture is already well represented in our model.

• Describe the negative trend of regional glacier mass balance for the WKSK in Figure 7(d).

Reply: Our conclusions on the mass balances are not based on the trend of the mass balances, but on the mean values, based on relative changes in temperature and precipitation. However, such a trend might indeed be interesting in saying something about the future. However, in this case, the trend is very insignificant, with a standard error more than twice the magnitude of the trend, and extrapolating to the future would not be meaningful. We describe the trend now in the discussion:

“**Our modelled mass balances show a decreasing mass balance trend for WKSK (Fig. 9d), but the trend is far too insignificant to draw conclusions about future mass balances.**”

• Section 3.1: Within the upper Indus Basin, observations from a number of automated weather stations are available from SIHP, Pakistan since 1994/1995 up to 4440 m asl and from the long-term PMD stations up to 2200 masl since the 1960s or earlier. For example in Norris et al., 2018. Importantly, SIHP automated weather stations include both snow and rain. Additionally, snow heights and snow fall amounts from the Weather Monitoring Posts from the Agha Khan Agency for Habitat and a few observations from EVK2-CNR are a valuable database for validation. For this, station selection criteria can be further relaxed to the available observations as the validation of the whole length of simulation does not seem to be mandatory in the data scarce region.

Reply: The stations used by Norris et al. and many other Pakistani stations are not easily accessible, do not cover the region with glaciers that have positive mass balances, and are mainly located on valley floors. Furthermore, even when snow measurements are present, they are highly uncertain. Therefore, we think they are unlikely to greatly increase our confidence in our modelling of snow in WKSK. We have now added comparison with the less uncertain AIRS and AMSU retrievals to further give confidence in our modelling of the moisture transport (see above).

• Lines 399-400: Unlike Waqas and Athar (2018), several studies suggest statistically significant strong cooling at least in July and September months over both low and high-altitude stations within the HKH region.

Reply: When making such comparisons, it is important to compare similar time periods. Indeed, we also show recent cooling in different datasets, including our own, in Fig. 3 for a recent time period. However, it is clear from Fig. 8 that over the longer time-period, there is nevertheless an overall warming, despite this recent cooling. There is also evidence for a cooling trend before 1980 (e.g. the work by Fowler and Archer, 2006), but this is not considered in our study. We now add: “**The temperature increase is there despite a recent decrease in summer temperatures in the region (Fig. 3).”**

• The study mainly focuses on reproducing the annual mass balances of the WKSK glaciers featuring delicate changes using a highly simplified lumped mass balance model. It would have been better to model the mass balance of these glaciers using a more sophisticated model and on an intra-annual scale as measurements/variables from the model are not an issue.

Reply: Such simplicity is needed to be able to model the tens of thousands of glaciers in the region within a reasonable time. We add to the Methods: “**The aforementioned model is relatively simple, but such simplicity is required to model the thousands of glaciers in the region within a reasonable time. Other models that aim to model glacier mass balances over such a large scale are also relatively simple, and the output of our model is close to the median of similar models for High Mountain Asia, with our model being the only one that treats debris cover (Marzeion et al., 2020)..”**

More detailed modelling is certainly encouraged in future work, but, given the big step in understanding the Karakoram Anomaly we already make in this paper, it is beyond the scope of the current paper. We add to the discussion, when discussion more detailed studies of smaller regions: “**It is then also possible to use more complex glacier models, e.g. those that take into account the full energy balance.”**

Towards understanding the pattern of glacier mass balances in High Mountain Asia using regional climatic modelling

Remco J. de Kok[1], Philip D.A. Kraaijenbrink[1], Obbe A. Tuinenburg[2], Pleun N.J. Bonekamp[1], Walter W. Immerzeel[1]

[1]Department of Physical Geography, Utrecht University, Utrecht, PO Box 80115, 3508 TC, The Netherlands
[2]Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, PO Box 80115, 3508 TC, The Netherlands

Correspondence to: Remco J. de Kok (r.j.dekok@uu.nl)

Abstract. Glaciers in High Mountain Asia provide an important water resource for communities downstream and they are markedly impacted by global warming, yet there is a lack in understanding of the observed glacier mass balances and their spatial variability. In particular, the glaciers in the western Kunlun Shan and Karakoram ranges (WKSK) show neutral to positive mass balances despite global warming. Using models of the regional climate and glacier mass balance, we reproduce the observed patterns of glacier mass balance in High Mountain Asia of the last decades **within uncertainties**. We show that low temperature sensitivities of glaciers and an increase in snowfall, for a large part caused by increases in evapotranspiration from irrigated agriculture, result in positive mass balances in WKSK. The pattern of mass balances in High Mountain Asia can thus be understood from the combination of changes in climatic forcing and glacier properties, with an important role for irrigated agriculture.

1. Introduction

Glaciers in High Mountain Asia (HMA, see Fig. 1) show a very diverse response to the changing climate. Most glaciers are losing mass, as expected, but surprisingly, glaciers are stable or growing in a region northwest of the Tibetan plateau, a phenomenon dubbed the Karakoram anomaly (Hewitt, 2005). Recent studies have mapped glacier mass losses and velocity changes in detail and have shown that the regions of largest glacier growth and acceleration are the Kunlun Shan and parts of the Pamir, with the glaciers in Karakoram being close to a steady state (Brun et al., 2017; Dehecq et al., 2019; Gardelle et al., 2012, 2013; Kääb et al., 2015; Lin et al., 2017). Part of the mass balance variability seems correlated to differences in the temperature sensitivity, i.e. the change of mass balance for a certain change in temperature, of the glaciers (Sakai and Fujita, 2017), but this alone cannot explain why some glaciers are actually growing, since either a decrease of ablation or an increase in accumulation is needed.

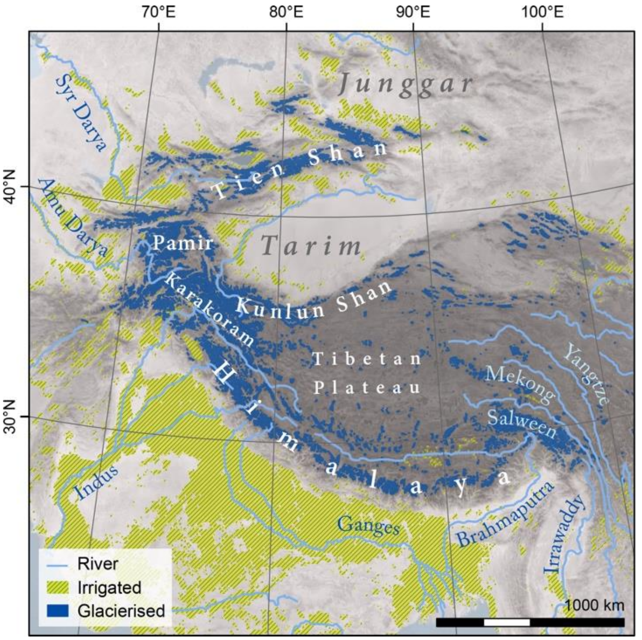


Figure 1: Map of study area. Irrigated areas (from GMIA (Siebert et al., 2010)) and glacierised areas (from RGI 6.0 (Pfeffer et al., 2014)) are indicated.

Suggested causes of the Karakoram anomaly include an increase in winter snowfall (Cannon et al., 2015; Kapnick et al., 2014; Norris et al., 2015, 2018), summertime cooling (Bocchiola and Diolaiuti, 2013; Forsythe et al., 2017; Fowler and Archer, 2006; Khattak et al., 2011; Ul Hasson et al., 2017), and an increase in summertime precipitation and clouds due to irrigation in the agricultural regions adjacent to the Kunlun Shan and Pamir (de Kok et al., 2018). So far, these hypotheses have only tried to explain the Karakoram anomaly in qualitative terms, identifying possible climatic conditions that could lead to glacier growth. None of these have yet been able to directly reproduce the observed pattern of glacier mass balances in HMA directly by showing the response of the glaciers to the historic climatic forcing. Here, we simulate glacier mass balances using modelled time series of temperature and snowfall from a regional climate model to reproduce the pattern of observed mass balances in HMA, and to more deeply understand the underlying causes.

2. Methods

2.1 Regional climate model

Irrigation can influence the regional and global climate (Cook et al., 2015; Lee et al., 2011; Lobell et al., 2008; Puma and Cook, 2010; Sacks et al., 2009). Since the regions surrounding HMA host the largest irrigated areas in the world, e.g. the Indo-Gangetic Plain (see Fig. 1), irrigation potentially influences the regional climate in HMA (Cai et al., 2019; de Kok et al., 2018).

However, available re-analysis datasets do not fully include irrigation, and generally have a relatively coarse spatial resolution. Hence, we downscaled ERA-Interim (Dee et al., 2011) re-analysis data using the Weather Research and Forecasting model (WRF, Skamarock & Klemp, 2008) to obtain a climate dataset between 1980-2010 for High Mountain Asia with a resolution higher than that of ERA-Interim. We artificially applied irrigation to the surface by adding a precipitation term each time step, with a rate that is determined per month. **The precipitation is added to the NOAH-MP surface module, but the atmospheric module is not altered.** The amount of irrigation applied per grid cell was based on the monthly water demand, which indicates the amount of irrigation needed to compensate evapotranspiration, after subtraction of the precipitation, that was calculated by the PCRaster Global Water Balance model (PCR-GLOBWB; van Beek & Bierkens, 2008; van Beek, Wada, & Bierkens, 2011; Van der Esch et al., 2017; Y. Wada, Wisser, & Bierkens, 2014; Yoshihide Wada et al., 2011). In this way, the irrigation amount is not easily overestimated, as could be the case when e.g. soil moisture would be kept constant. In reality, insufficient water might be available to meet the predicted demand, whereas inefficient irrigation will result in a larger water gift than predicted. **The PCR-GLOBWB runs were forced by WATCH data, based on ERA-Interim (Weedon et al., 2014).**

We used WRF, version 3.8.1, to downscale ERA-Interim data to a resolution of 20x20 km, with 50 vertical levels. Settings are nearly identical to our previous work (de Kok et al., 2018), **which are based on the work of Collier and Immerzeel (2015) and Bonekamp et al. (2018)** and are shown in Table 1. Additionally, we now use grid nudging of the upper 35 vertical levels for horizontal wind, temperature, and humidity, as opposed to only forcing the model at the boundary in our previous study. This ensures the large-scale upper-atmospheric circulation closely follows that of ERA-Interim, whereas near the surface, the model is more determined by the physics in WRF, e.g. evaporation in irrigated areas. The nudged levels and the values of the nudging parameters have been found to perform well in similar studies (Collier and Immerzeel, 2015; Otte et al., 2012), and are: 0.0001, 0.0001, and 0.00005 s^{-1} for horizontal winds, temperature and water vapour, respectively. The default values for all three parameters are 0.0003 s^{-1} in WRF.

Table 1: Physics modules and assumptions used in WRF.

Module	Setting
Radiation	RRTMG scheme (Iacono et al., 2008)
Microphysics	Morrison scheme (Morrison et al., 2009)
Cumulus	Kain-Fritsch (new Eta) scheme (Kain, 2004)
Planetary boundary layer	YSU scheme (Hong et al., 2006)
Atmospheric surface layer	MM5 Monin-Obukhov scheme (Beljaars, 1995; Dyer and Hicks, 1970; Paulson, 1970; Webb, 1970; Zhang and Anthes, 1982)
Land surface	Noah-MP (Niu et al., 2011)

Top boundary condition	Rayleigh damping
Diffusion	Calculated in physical space
Nudging	Grid-point *u*, *v*, *T*, *q* above level 15

Annual concentrations of CO_2, CH_4, and N_2O, which are manually set in the RRTMG radiation module, are derived from NOAA (Dlugokencky et al., 2018) and AGAGE (Prinn et al., 2000) data, as aggregated by the European Environment Agency (www.eea.europa.eu, accessed March 2018), and are kept constant throughout each year. We used a 10-day spin-up for each month and run each month separately to be able to include a different irrigation amount each month. Monthly initialisation of the snow cover, surface and soil temperature, and surface moisture was taken from GLDAS 2.0 (Rodell et al., 2004) monthly mean values. We checked whether temperatures and precipitation at the end of a month agreed with those at the end of the spin-up period for the subsequent month and they agreed within a few percent for all selected points. Results are output every 6 hours.

2.2 Glacier model

To assess the response of the glaciers to the atmospheric forcing, we employ a glacier mass balance gradient model (Kraaijenbrink et al., 2017). The model assumes a calibrated mass balance gradient along the glacier, and parameterises downslope mass flux in a lumped procedure that is based on vertical integration of Glen's flow law (Marshall et al., 2011). It also includes a parameterisation for the effects of supraglacial debris on surface mass balance (Kraaijenbrink et al., 2017), i.e. enhancing melt in the case of a shallow debris layer and limiting melt for thicker debris (östrem, 1959). We modelled all individual glaciers in HMA larger than 0.4 km^2 (n=33,587) transiently for the period 1980-2010 (Kraaijenbrink et al., 2017). For ease of comparison with published observations, we select only those larger than 2 km^2 for the final analysis, which represent 95% of the glacier volume in HMA. Initial mass balance conditions in 1980 were set to be stable, while all other initial and reference conditions as described in the original study (Kraaijenbrink et al., 2017) were maintained. That is, using ERA-Interim data to locally calibrate the mass balance gradient of each glacier by constraining maximum ablation by a downscaled positive degree day climatology at the glacier terminus, and maximum accumulation by mean annual precipitation over the entire glacier area. The model simulates glacier mass change and evolution using a one-year time step, and hence requires representative annual input of temperature and precipitation. These are used to shift the mass balance curve according to sensitivity of the glacier's equilibrium line altitude to temperature changes, and adapt the maximum accumulation according to changes in precipitation (Kraaijenbrink et al., 2017). **The aforementioned model is relatively simple, but such simplicity is required to model the thousands of glaciers in the region within a reasonable time. Other models that aim to model glacier mass balances over such a large scale are also relatively simple, and the output of our model is close to the median of similar models for High Mountain Asia, with our model being the only one that treats debris cover (Marzeion et al., 2020).**

To modulate the curve transiently, we applied annual precipitation changes derived from annual changes in WRF snowfall and **air** temperature changes determined from annual changes in WRF melt season temperatures, i.e. when average daily **air** temperature is above -5 °C. A threshold value of 0 °C did not significantly change the glacier mass balance results for most of HMA, but meant that temperature changes for the coldest points could not be determined, since they are always below 0 °C. We did not take into account whether the WRF grid point was glacierised or not. To reduce potential biases imposed by spurious reference conditions, the reference for the changes in **air** temperature and precipitation was taken to be the average of the first six modelling years, i.e. 1980-1985. We performed three separate glacier model runs to evaluate the attribution of snow**fall** and temperature to the glacier mass changes in our model, forced by: (1) precipitation and **air** temperature, (2) only **air** temperature, and (3) only snow**fall**. To illustrate the temperature and precipitation sensitivity of the glaciers, we also performed calculations using a fixed **air** temperature or snowfall trend for all of HMA, with the other variable kept constant.

2.3 Moisture tracking

Our moisture tracking model (Tuinenburg et al., 2012) follows the moisture associated with precipitation backwards in time to determine where the moisture first enters the atmosphere. It therefore establishes a direct causal link between evapotranspiration and precipitation downwind. For the moisture tracking, we clustered locations that have similar climates in terms of seasonality, since these will likely also have similar moisture sources. For the clustering, we used the monthly mean values of precipitation, horizontal wind fields at 400 hPa, and 2m-temperatures, with means subtracted and divided by the standard deviations, to perform a k-means clustering using 13 clusters. In this way, we delineate regions that have similar surface variables, relevant for the glaciers. Furthermore, these regions are also under the influence of similar winds, relevant for the moisture transport. We performed the clustering with different numbers of clusters and found 13 to give reasonably-sized, yet distinct areas, while also being close to the knee in the total distance away from the cluster means. We perform the tracking on two of these clusters, indicated later in Fig. 14, which are close geographically, but have contrasting snowfall trends.

We perform the moisture tracking by releasing moisture parcels from the target area at random positions within the area and advecting them backwards in time using interpolated wind fields on fixed pressure levels. The amount of evaporation A (mm) that contributes to the precipitation in the target area, at a given location x,y and time step t, depends on the evapotranspiration ET (mm), the **total** amount of **all** water in the parcel W_{parcel} (mm), the fraction of water in the parcel that evaporated from the source S_{target}, and the total precipitable water in the column TPW (mm):

$$A_{x,y,t} = \mathrm{ET}_{x,y,t} \frac{W_{parcel,t} S_{target,t}}{\mathrm{TPW}_{x,y,t}} \quad , \qquad (1)$$

The amount of water in the parcel is then updated every time step, including the precipitation *P* that adds to the parcel when moving back in time.

$$W_{parcel,t-1} = W_{parcel,t} + \left(P_{x,y,t-1} - \mathrm{ET}_{x,y,t-1}\right)\frac{W_{parcel,t}}{TPW_{x,y,t-1}} \quad (2)$$

The fraction **of** precipitation in the target area that originates from a certain source area is then updated as follows:

$$S_{target,t-1} = \frac{S_{target,t}W_{parcel,t} - A_{x,y,t-1}}{W_{parcel,t-1}} \quad (3)$$

We track the parcel until either more than 99% of the target precipitation is tracked to a source area, or the tracking time is more than 30 days.

Within the WRF domain, the parcels are advected and the moisture budget is calculated using the WRF wind fields and water fluxes. When a parcel gets within one degree of the edge of the WRF domain, there is a gradual linear change to a use of ERA-Interim reanalysis data to ensure continuous movement of the air parcels over the domain edge. Within one degree distance from the domain edge, the values used to do the moisture tracking are a combination of the WRF and ERA-Interim values: $y_{int} = d*y_{WRF} + (1-d)*y_{ERA}$, where *d* is the distance to the edge. Outside of the WRF domain, the ERA-Interim values are used.

We noted that the surface moisture flux in ERA-Interim is on average 50% higher than in WRF, resulting in a higher mean and standard deviation of the moisture sources outside the WRF domain. Unfortunately, this systematic offset between the two datasets cannot be easily remediated. Although this will not change the spatial patterns of the moisture source trends in a major way, the absolute values of the trends will be lower in the WRF domain than outside if there is a scaling factor in moisture flux between the two datasets. The trends in the Tarim basin will then be underestimated with respect to regions such as the Caspian Sea and the Junggar basin.

2.4 Statistics

Pearson correlation coefficients are calculated **between pairs of different datasets (e.g. Figs. 2-5)** using the vectors of annual or seasonal mean values, **with one value for each year**. **The figures indicate over which period the mean is taken for each year.** The trends shown in Figs. 2-5, 8, and 17 are the slopes from linear fits to **these** vectors. **P-values for the correlations are determined using the beta function, as implemented in SciPy (Virtanen et al., 2020).**

3. Results

3.1 Validation and comparison

Any attempt to understand the Karakoram anomaly is greatly hampered by the almost complete lack of *in situ* meteorological data in WKSK. The sparse weather stations in the region are often situated at relatively low elevation, or in urban environments, and poorly represent the high mountain climate. Furthermore, different types of precipitation datasets seem to greatly underestimate the precipitation in mountainous terrain (Immerzeel et al., 2015; Ménégoz et al., 2013; Palazzi et al., 2013). These complications imply that any meteorological dataset, including reanalysis datasets, are associated with relatively large fundamental errors in WKSK, which prevents reliable validation of any model of WKSK, such as the one presented in here. Although not covering the glacierised areas of interest, we compared our WRF output with data of the region surrounding WKSK, to ensure that the WRF output is a reasonable representation of the regional climate between 1980-2010. Since the glacier model requires annual input, representation of the interannual variability is especially important. Any constant biases are of less importance, since we use relative interannual variations as input for the glacier model. However, biases in temperature will have an effect on the snow-rain partition.

We collected meteorological station data from the Global Historical Climatology Network (GHCN, Lawrimore et al., 2011, accessed June 2019), and selected those that have at least 15 years of full data between 1980-2010. To be able to compare the WRF output with the station data, we apply a simple downscaling to the WRF temperatures in the grid that includes the station. We fit a linear temperature lapse rate to the temperatures and grid altitudes of a 2x2° box surrounding the station location. We then correct the WRF temperature by applying the lapse rate to the difference in altitude between the WRF grid and the station. Precipitation can also change significantly with location, but there is no clear relation between precipitation and altitude (Bonekamp et al., 2019; Collier and Immerzeel, 2015). For this simple comparison, we do not apply a downscaling of the WRF precipitation.

Our WRF output produces May-September temperatures that are generally higher than the stations in the Tarim basin. However, biases are generally very low on the Tibetan Plateau, with values around 1°C **(Fig. 2a)**. The median root-mean-square deviation between WRF and the stations **for the time-series of seasonal mean temperatures** is 1.8°C. The stations generally indicate a strong heating trend **(Fig. 2b)**, **but also show relatively large differences for close-by stations**. Correlations between the annual variations in annual mean temperatures and mean temperatures between May-September are given in Fig. 2. They show generally very high correlations, with a lowest value of 0.5 (corresponding to $p = 0.005$, **Fig. 2c**). This implies that the interannual variability is very well reproduced in WRF. This is despite the fact that many of these stations are situated in urban environments, with a potential heat island effect, a lack of evaporative cooling that is seen for irrigated agriculture, and a very difference surface energy balance than snow-covered areas. Hence, their locations might not be representative of the wider area, which might give rise to biases and trend differences when comparing the stations to the model outcome.

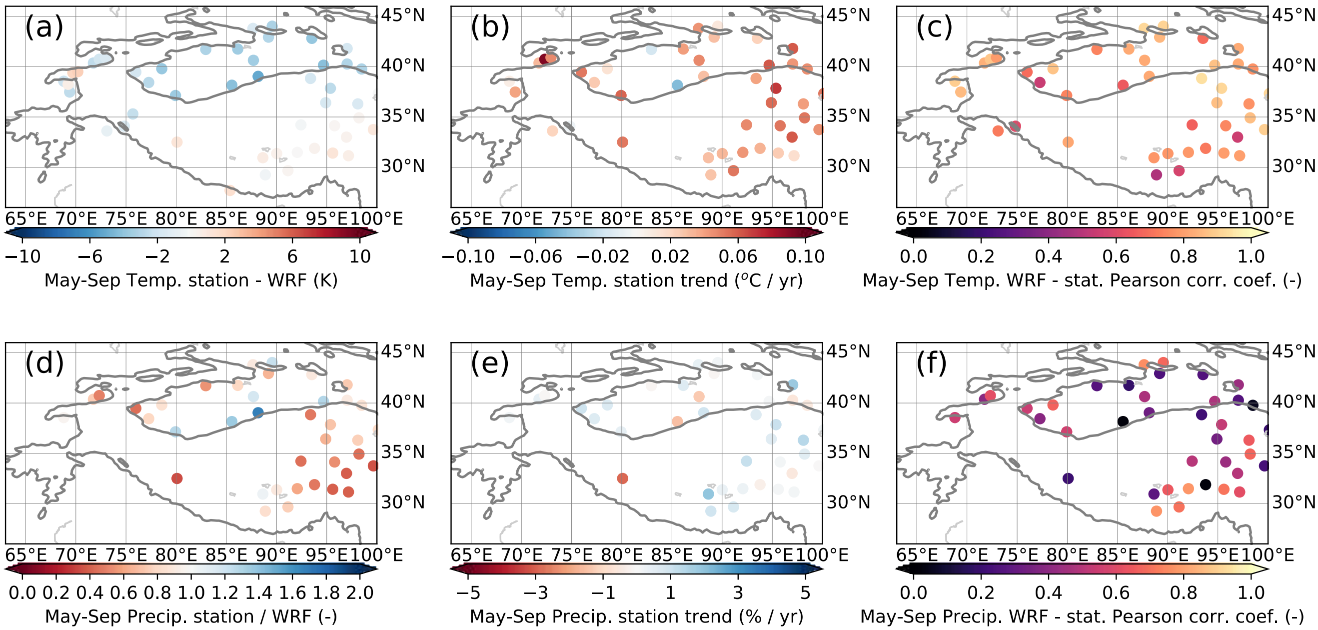


Figure 2: Comparisons between 1980-2010 time series of station data and nearest WRF grid point **for May-September temperatures (a-c) and May-September precipitation (d-f). Columns show temperature bias (a) and precipitation multiplication factor (d),** station **trends (b,e) and Pearson correlation coefficients. The 2000 m-contour is indicated by a solid line**

The stations in Fig. 2 closest to WKSK are almost exclusively in very arid regions, with a significant fraction of snowfall, which is more difficult to reliably measure than rain (Archer, 1998), making comparisons of precipitation very uncertain. Fig. 2 shows the comparison between time series of May-September precipitation, to limit the effect of snowfall. **The stations show an increasing trend in May-September precipitation in the western Tarim basin and most of the eastern Tibetan Plateau (Fig. 2e).** Our WRF output is generally wetter than what is measured at the stations, except some locations in the Tarim basin **(Fig. 2d)**. The median root-mean-square deviation between WRF and the stations **for the time-series of seasonal mean precipitation** is 11.4 mm per month. The stations show that most of the Tarim basin and Tibetan Plateau are seeing an increase in May-September precipitation. The interannual variations are not represented by WRF as well as they are for temperature, but still show reasonable correlations for most stations, with values around 0.6 **(Fig. 2f)**.

We also compare our WRF simulations with three similar data products with relatively high spatial resolutions, which have recently become available. We do note that all these datasets suffer from the lack of ground truth in WKSK, which means we cannot determine which dataset performs best in this region.

ERA5 is the follow-up of ERA-Interim (Copernicus Climate Change Service, 2017), with an improved spatial resolution of 0.25°, an improved temporal resolution, a more appropriate model input for e.g. sea surface temperatures, and more assimilated data. ERA5-Land is atmospherically forced by ERA5, and provides an even higher spatial resolution (0.1°) for land surface properties (Anon, 2019). Finally, we include the HAR dataset with a resolution of 10 x 10 km, which uses WRF to downscale the NCEP FNR reanalysis dataset and re-initialises every day (Maussion et al., 2014). We compare temperatures between May-September, and annual precipitation, which give an indication of the parameters that are most relevant for glacier mass balance modelling. Because of the limited time overlap between the different datasets, we could only fully compare the period 2001-2010.

We binned all data to the same 0.5° x 0.5° grid to allow direct comparison. The mean values, trends, and interannual variability are compared in Figs. 3 and 4. It shows that ERA5 and ERA5-Land are nearly identical, and we only refer to ERA5 below. Our WRF model yields a warmer Karakoram than the other three datasets. Generally, the mean temperature differences are relatively minor, except for a warmer Tarim basin compared to HAR. We find very similar temperature trends as ERA5, although with smaller magnitudes. The magnitudes of the trends are also generally smaller than those in the station data (Fig. 2**b**). The WRF interannual temperature variations correlate very well with ERA5, except two areas in the Tarim and the inner Tibetan Plateau. This is not surprising, given that our WRF model is forced by the similar ERA-Interim data. The whole western part of HMA, including WKSK, is especially well-correlated to ERA5. In that region, the correlation with HAR is weaker, but the correlation between HAR and our WRF data is very strong in East HMA. The differences with HAR might be explained by the different forcing, or by the difference in used physics modules, but this requires further study.

Differences between datasets are larger for precipitation, at least for the mean values and interannual variability. Our WRF simulations give results that are relatively wet in the Karakoram, and relatively dry in the Himalaya. However, the precipitation trends are very similar to ERA5 in both pattern and magnitude. An exception is the arid Tarim basin, which has an increasing trend in WRF, but a decreasing trend in ERA5. HAR shows a positive precipitation trend in most of HMA, with a very high trend in the Tarim basin. The correlation of the interannual variability is low in WKSK and parts of Tien Shan, which could be explained by the relatively high influence of the irrigated areas in the Tarim basin on the annual precipitation (**Fig. 3 of** de Kok et al., 2018). Since our WRF model outcome is the only one of the four datasets that explicitly includes irrigation, this could explain the difference in annual variability.

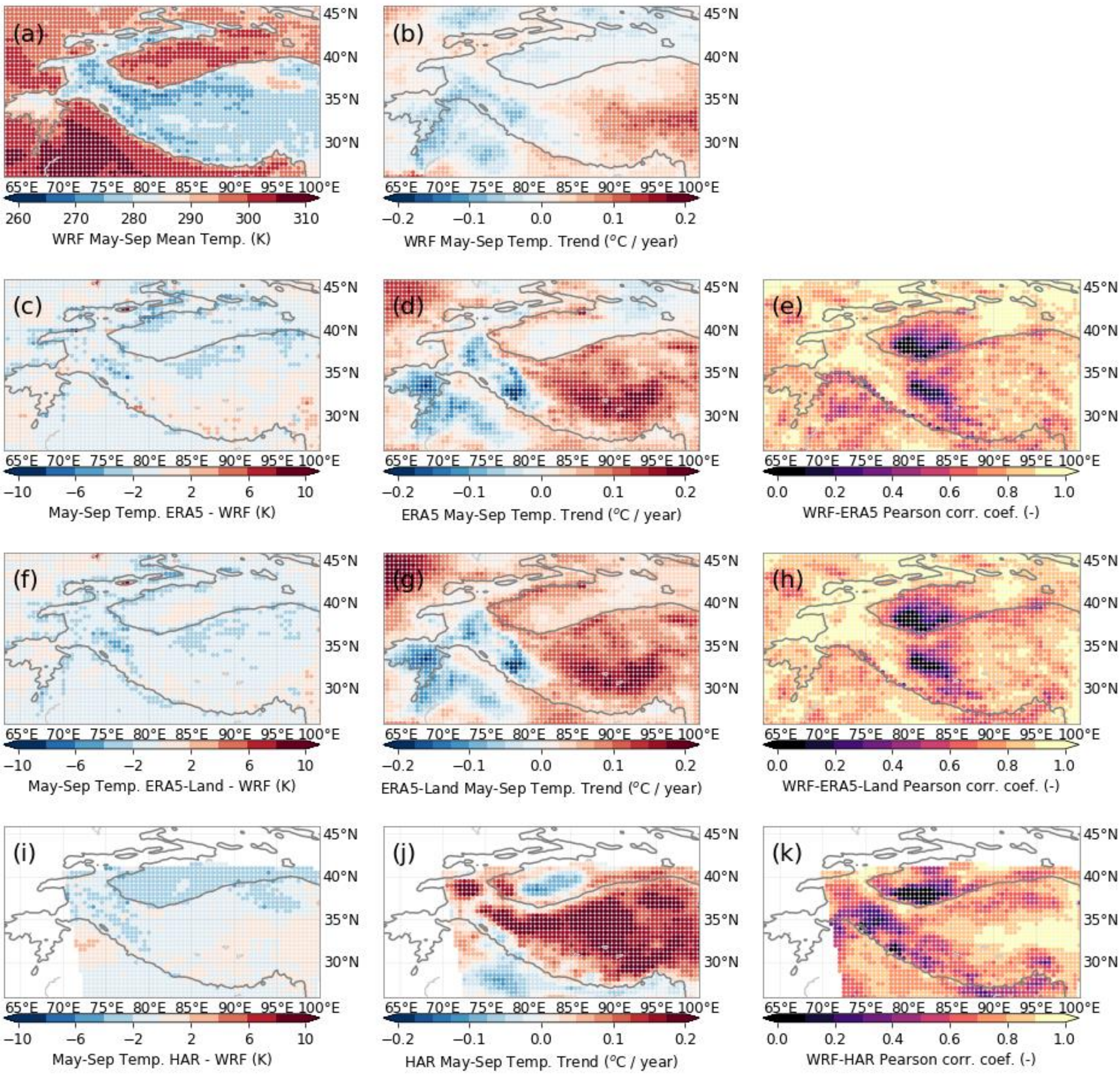


Figure 3: Comparison of WRF temperature output [a-b] with three other datasets (ERA5 [c-e], ERA5-Land [f-h], and HAR [i-k]). Columns show biases (c,f,i) with respect to the May-September mean temperature (a), May-September temperature trends (b,d,g,j), and Pearson correlation coefficients between the datasets and our WRF results (e,h,k). The 2000 m elevation contour is indicated by a solid line.

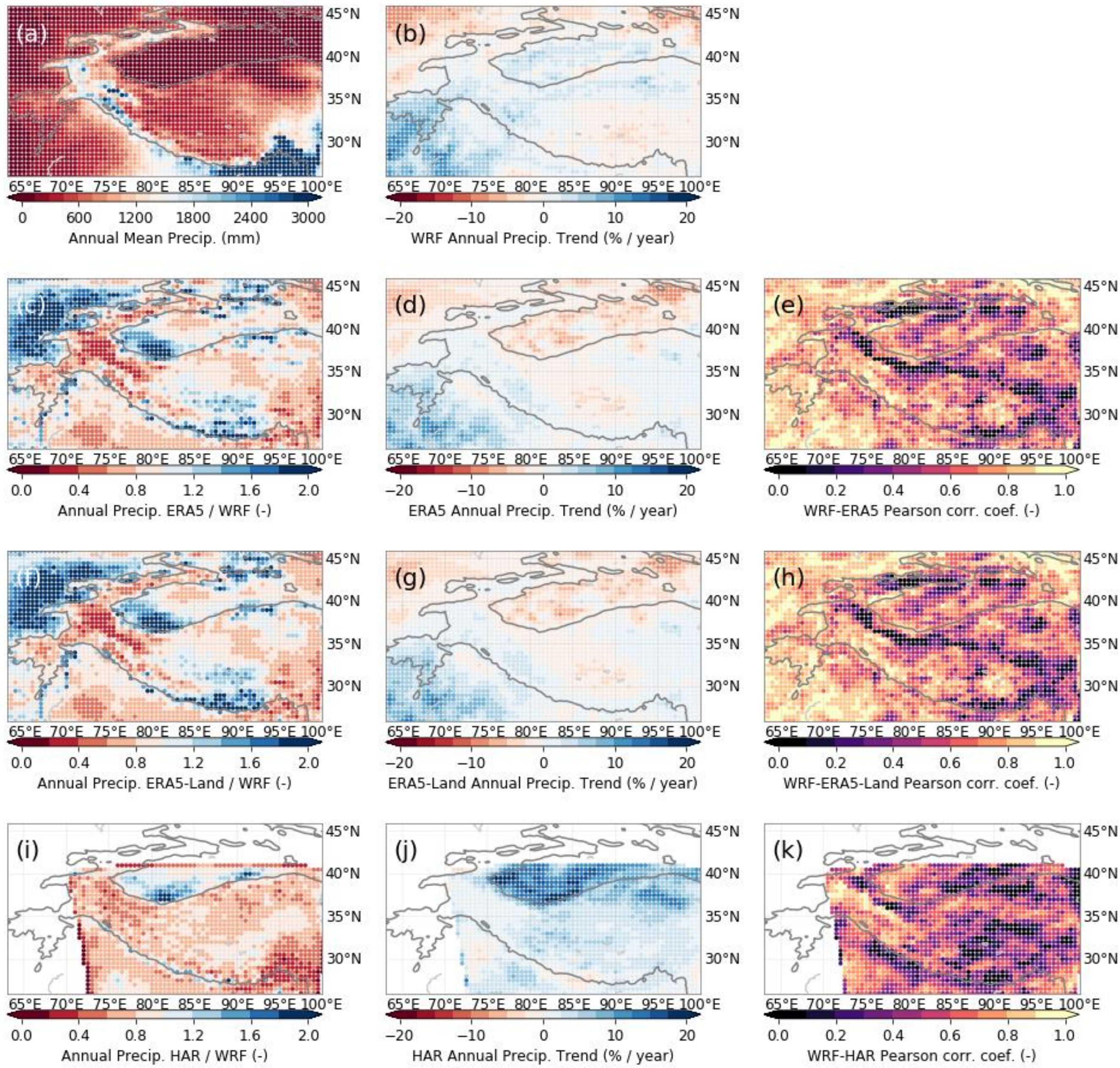


Figure 4: Comparison of WRF precipitation output [a-b] with three other datasets (ERA5 [c-e], ERA5-Land [f-h], and HAR [i-k]). Columns show precipitation multiplication factors (c,f,i) with respect to the annual mean precipitation (a), annual precipitation trends (b,d,g,j), and Pearson correlation coefficients between the datasets and our WRF results (e,h,k). The 2000 m elevation contour is indicated by a solid line.

The station and reanalysis data show a good agreement with our WRF output in many locations, but the comparison is hampered in WKSK due to the aforementioned fundamental uncertainties. Remote sensing data can also be used for comparison, but also there, uncertainties can be very high. This is especially true for precipitation measurements in mountainous areas, but also other remote surface measurements of relevant parameters are uncertain in mountainous areas (Lundquist et al., 2019). However, the atmosphere above the mountains can be measured with some confidence. Especially the atmospheric humidity can be used to increase the confidence in the interannual variability of the precipitation, since the two are strongly related. Here, we compare retrieved atmospheric humidity from AIRS and AMSU data (AIRS Science Team and Teixeira, 2013) above the mountains with humidity from our WRF output. These retrievals determine the humidity from satellite measurements at wavelengths in the infrared and microwave with very limited assumptions (Susskind et al., 2014), and hence can be considered as good validation dataset. Figure 5 shows the comparison of the mean AIRS specific humidity between May-September and between 400-500 hPa, and the corresponding WRF specific humidity interpolated at the middle of this layer (447.2 hPa) for the overlapping years 2003-2010, binned at the AIRS-AMSU resolution. The two datasets show a very high overall agreement, both in the patterns of humidity trends, as well as the correlation of interannual variability, with correlation coefficients generally above 0.9. This analysis shows that the moisture transport in our WRF model closely follows what we know of the atmosphere around WKSK. Near the edges of our modelling domain, our errors are naturally larger.

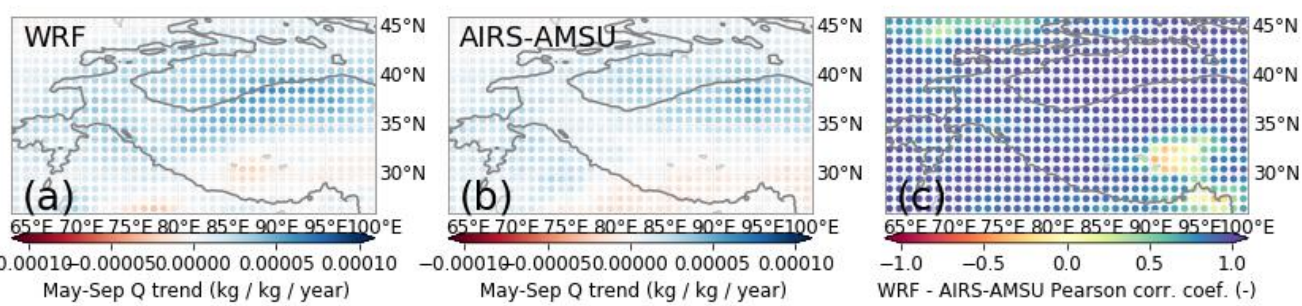


Figure 5: May-September mean specific humidity trends at 447.2 hPa for WRF (a), and between 400-500 hPa for AIRS-AMSU (b) between 2003-2010. Panel (c) shows the Pearson correlation coefficient between them.

Another variable that is important in our model is evapotranspiration. It cannot be directly measured remotely, but there are several datasets that calculate it from other remotely sensed products, either directly or through data assimilation. These datasets **are all validated to some extent, but** vary greatly **nevertheless**, as we illustrate in Fig. 6 for July 2010. We show evapotranspiration from GLEAM v 3.3a (Martens et al., 2017; Miralles et al., 2010), which assimilates various soil moisture, temperature, radiation, and precipitation products. Furthermore, we show SSEBop (Senay, 2018) data, which uses MODIS temperatures directly, ERA-Interim reanalysis data, and our WRF output. On the inner Tibetan Plateau, the WRF output agrees very well with the GLEAM data. Interannual variations also match very well between WRF and GLEAM in snow-free areas on the Tibetan Plateau, with correlation coefficients above 0.5 for time series between 1980-2010. However, it is clear that GLEAM does not represent the irrigated areas well, with evapotranspiration in heavily irrigated arid regions in July **that is as low as the surrounding deserts in e.g. Tarim and Indus basins, which is not realistic**. In contrast, SSEBop shows very high evapotranspiration in the irrigated regions. The WRF output better resembles SSEBOP in those areas, although generally has lower maxima, which are only in part explained by the difference in spatial resolution, as is evident from e.g. averaging over 1x1° areas. ERA-Interim does not show the irrigation as prominently as WRF or SSEBop, but has a generally higher evapotranspiration values over unirrigated areas, such as the Tibetan Plateau. In general, the WRF simulated evapotranspiration is intermediate compared to the other datasets with plausible spatial patterns and magnitudes. **However, the figure illustrates the problem with the high uncertainty in evapotranspiration over large areas in and around HMA.**

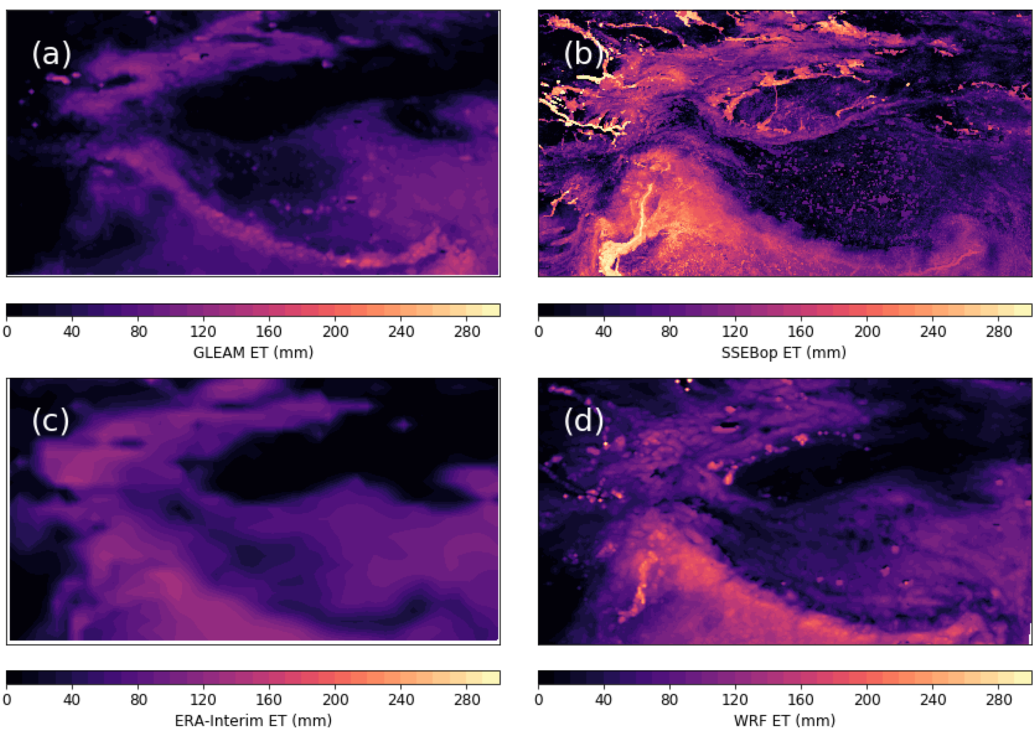


Figure 6: Evapotranspiration for July 2010 from GLEAM (a), SSEBop (b), ERA-Interim (c), and WRF (d). Mean values for the plotted domains are: 44 mm (a), 46 mm (b), 57 mm (c), and 59 mm (d).

We further investigate the realism of the effect of irrigation in our model by comparing remotely sensed surface specific humidity from AIRS and AMSU retrieval with our WRF specific humidity at 2 metres. These are not exactly the same quantities, as AIRS has a finite vertical resolution, but the variations over time can be compared. We focus on the irrigated area in the Tarim basin, close to the Kunlun Shan, which is the most important in the later discussion on the Karakoram anomaly. The flat terrain makes the retrievals near the surface more certain compared to mountainous regions, where altitude, and hence pressure and humidity, strongly vary within the spatial resolution of the measurements. The comparison between means over May-September for 2003-2010 is shown in Fig. 7. Even though we did not nudge WRF towards ERA-Interim near the surface, the model still follows the humidity observations in the irrigated region in the Tarim basin very closely, with a Pearson correlation coefficient of 0.97. This gives further confidence that the irrigation we apply there is not unrealistic.

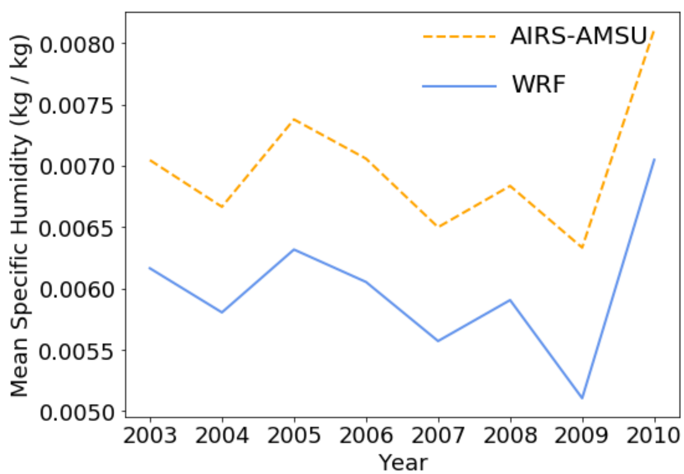


Figure 7: May-September mean specific humidity at 2 metres from WRF (blue, solid line) and AIRS-AMSU surface humidity (orange, dashed line) for a 1° x 1° bin around 38.5° N, 77.5° E, which is an irrigated area in the Tarim that contributes to the snowfall in WKSK.

3.2 Climatic trends

To get an impression how glaciers might have been affected by changes in the climate, we illustrate the trends for two relevant variables: the 2m-temperature in the melt season and the annual snowfall (Fig. 8, see also Fig. 9 for representative time series). For each grid point, the melt season was defined as the months where the mean daily temperature is above -5 °C, since for these months temperatures will likely be above freezing at least part of the time. A threshold value of 0 °C slightly increased the positive temperature trends at lower elevations in WKSK, but meant no trends for the highest elevations could be determined. The trends show that temperatures in the melt season have generally increased, with the northern part of the domain heating up the fastest and parts of the Indo-Gangetic Plain, Kunlun Shan, Karakoram, and the Tibetan Plateau showing only modest increases in temperature. **The temperature increase is there despite a recent decrease in summer**

temperatures in the region (Fig. 3). Fig. 9 shows that the trend and the interannual variability of temperature are very similar for nearby regions of both growing and shrinking glaciers. The snowfall trends in Fig. 8 have a very different pattern, with most of the Tibetan Plateau showing an increase and the western and southern mountain ranges, such as the Himalaya and the Hindu-Kush, showing a decrease in snowfall. Furthermore, the mean level, the trend, and the interannual variability of snowfall is quite distinct for the two nearby regions of contrasting glacier mass balance trends. The increase in snowfall in WKSK mainly occurs in May, June, and September, whereas the decrease of snowfall in southwestern HMA occurs mainly in March (see Fig. 10**d** for region averages).

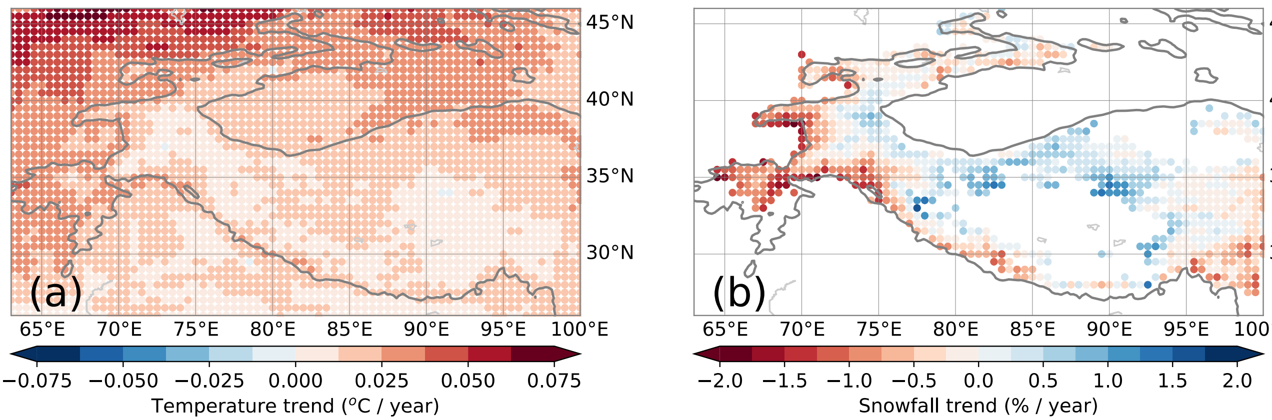
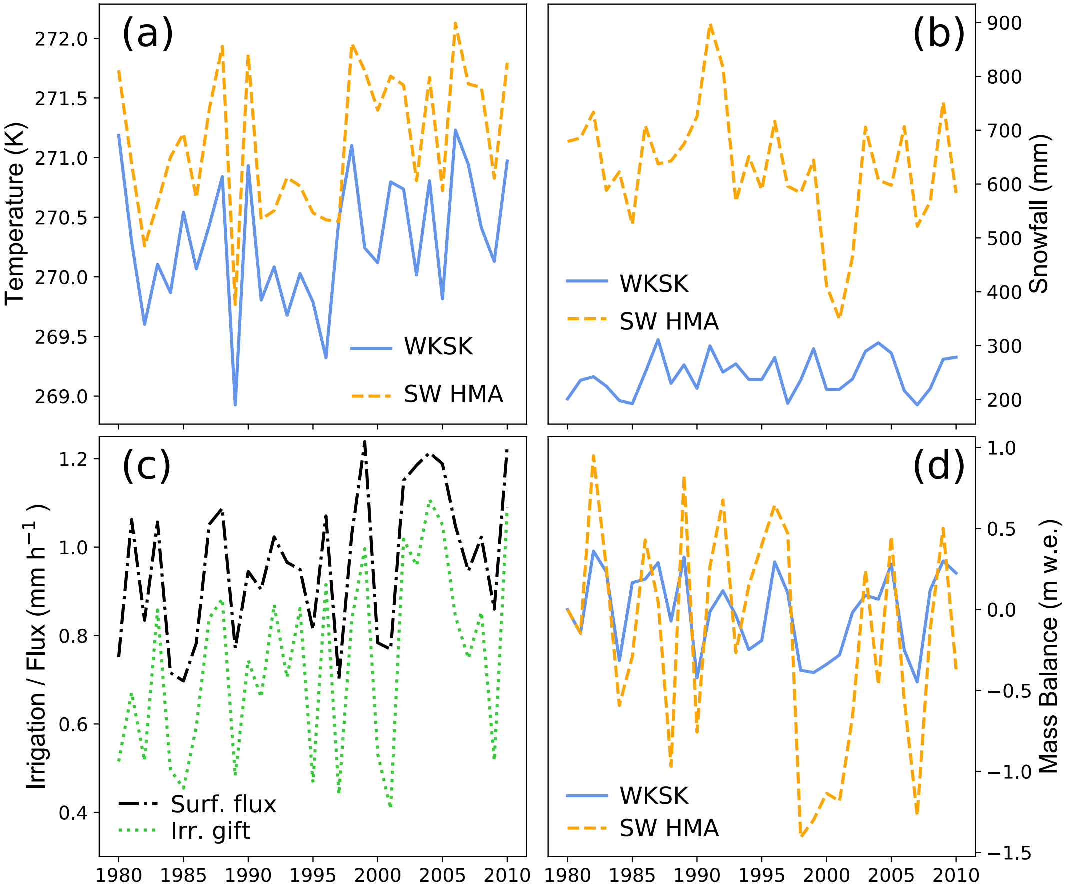


Figure 8: Trends between 1980-2010 of temperature in the melt season (a) and annual **snowfall (b), averaged over 0.5x0.5° bins for clarity. Regions with monthly snowfall of less than 10 mm were masked out. The 2000 m elevation contour is indicated by a solid line.**



340 **Figure 9: Time series of annual mean temperature (a), annual snowfall (b), mass balance (d) for two nearby 2x3° bins** in WKSK and southwestern HMA **that have, on average, growing glaciers (38-40° N, 73-76° E, blue lines) and shrinking glaciers (35-37° N, 72-75° E, orange, dashed lines). Panel *c* shows the time series of annual irrigation gift (green, dotted line) and annual surface moisture flux (black, dot-dashed line) for the most heavily irrigated point in the Tarim.**

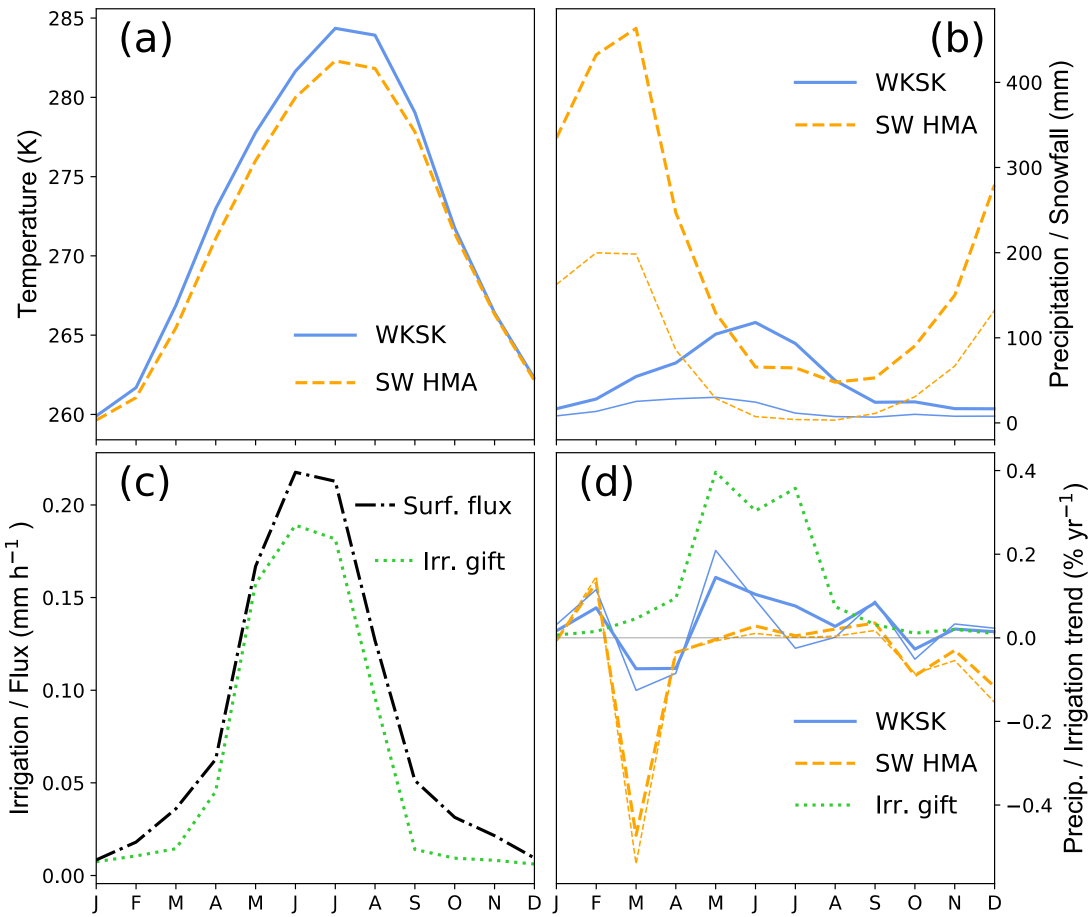


Figure 10: a) Mean seasonal cycle of temperature and b) precipitation (thick lines) and snowfall (thin lines) between 1980-2010 for the WKSK (blue lines) and southwestern HMA (orange, dashed lines), as shown in Fig. 14b,d. c) Mean seasonal cycle of the irrigation gift (green, dotted line) and surface moisture flux (black, dot-dashed line) of the most heavily irrigated point in the Tarim. d) Trends in precipitation (thick lines) and snowfall (thin lines) for the WKSK (blue lines) and southwestern HMA (orange, dashed lines), and the trend in irrigation from the most heavily irrigated point in the Tarim (green, dotted line), all as percentages of the annual mean value.

3.3 Glacier mass balances

The resulting pattern of simulated mass balance (Fig. 11) shows a strong resemblance to the measured pattern of mass balances of recent decades. Most notably, we also obtain growing glaciers in WKSK, whereas the glaciers in other regions show large mass losses. In fact, all points where we model glacier growth in Fig. 11a also show growth or stable conditions in observations (Brun et al., 2017; Kääb et al., 2015), except one point in Kääb et al., (2015). A more detailed quantitative comparison of the above results and the observed mass balances is hampered by the fact that our simulations only go out to 2010, and hence we cannot compare with the most recent, and most accurate geodetic mass balance data. However, we compare our results for the intermediate period 2000-2008, as presented by Brun et al. (2017), in Fig. 12. The results generally match reasonably well, although our model seems to show too little growth for the growing glaciers. However, note that the errors on these observations (Brun et al., 2017) are large (~0.3 m w.e.). Furthermore, both the climate model and the glacier model will be associated with errors. However, in both cases the growing glaciers are only present in the same region, mainly WKSK and the Tibetan Plateau. By modulating the initial mass balance in the model, we find that on average 41% of the modelled mass balance in 2010 is determined by the initial mass balance in 1980. Although the mass balances in 1980 were observed to be less extreme than in the 21st century (Bolch et al., 2012; Maurer et al., 2019), parts of HMA already had negative mass balances then, with the magnitude of initial mass balances generally less than 0.4 m w.e. yr^{-1}. This would result in an error on the mass balances in Fig. 10 of less than 0.2 m w.e. yr^{-1}. Despite these uncertainties, our results clearly show that the climatic change of the recent decades has favoured growth of the glaciers in the regions where actual growth is observed, and not in the places where glaciers are melting fastest.

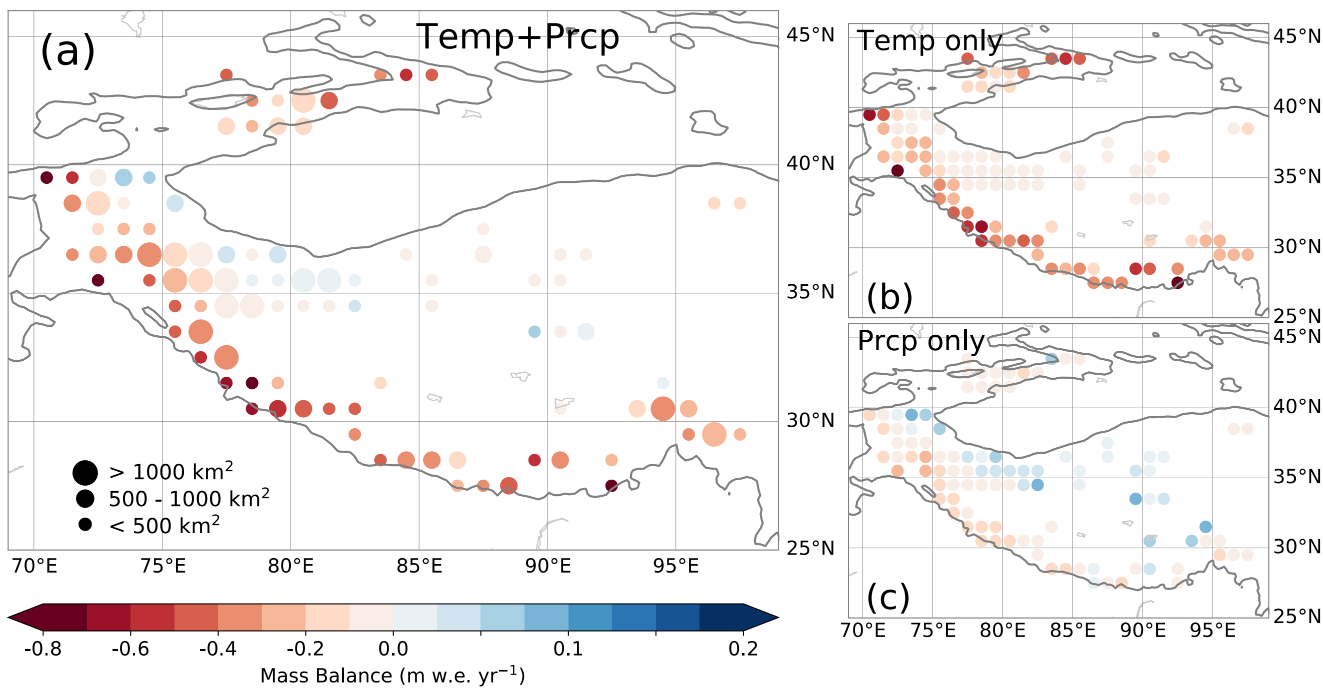


Figure 11: Simulated mean mass balance between 2000-2010 forced by changes in temperature and annual snowfall from WRF (a), only changes in temperature from WRF (b), and only changes in annual snowfall from WRF (c). Results are binned in 1x1° bins, and bins with total glacier volumes less than 5 km^3 are not shown, to enable comparison with previous studies. The 2000 m elevation contour is indicated by a solid line.

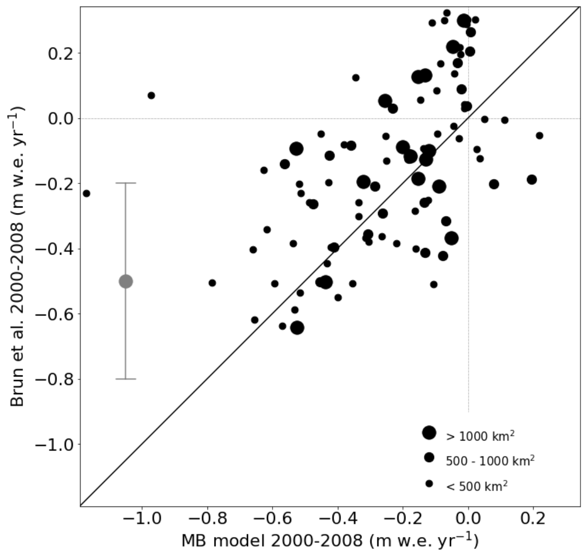


Figure 12: Comparison between mean modelled mass balances from this work, binned on a 1x1° grid as in Fig. 11, and those derived from observations of Brun et al. (2017), which are on the same grid, between 2000-2008. The size of the mean errors on the observed mass balances is illustrated by the grey error bar.

We also ran the glacier mass balance model forced by changes in temperature or snowfall only, to disentangle the model sensitivities of the two different variables on the glacier mass balances (Figures 11b and 11c). These results show that the glaciers in the western and southern HMA mainly lose mass due to the increase in temperature, while the decrease in precipitation gives a much smaller mass balance response in this region. On the other hand, in the regions where the glaciers are growing, the glaciers are barely affected by the temperature increases in our model. The glacier growth in these regions is mainly caused by an increase in snow**fall** (Fig. 11c). Furthermore, the increase in snow is possibly also responsible for moderating the temperature increases due to the high albedo of fresh snow, which leads to less energy being used for melt. However, the weak temperature response in WKSK is not only caused by the limited temperature trends, but is also due to the limited glacier temperature sensitivity there. We demonstrate this by forcing the glacier model with uniform temperature and precipitation trends (Fig. 13). The reduced temperature sensitivity is in line with previous work (Sakai and Fujita, 2017; Wang, et al., 2019), which argue that the generally large masses of the glaciers, and high equilibrium line altitudes, are important in explaining the lower temperature sensitivity in WKSK. The decrease in snowfall in the western and southern HMA has a far smaller impact on the mass balance than the increase in temperature. Especially the Himalaya show a low sensitivity to precipitation (Fig. 13). To be able to model thousands of glaciers, our mass balance model is relatively simple and does not solve the full energy balance. A full energy balance model at 1 km resolution has shown that the temperature increases can amplify melt in the monsoon-dominated Himalaya, whereas snowfall increases in the melt season can amplify glacier growth in the Karakoram (Bonekamp et al., 2019). Hence, more detailed models will likely strengthen our conclusion that the observed mass gains are caused by snow increases, whereas the observed mass losses are mainly caused by temperature increases.

Unfortunately, modelling the climate and glaciers of the entire HMA at a sub-kilometre resolution for 30 years is currently beyond our capabilities.

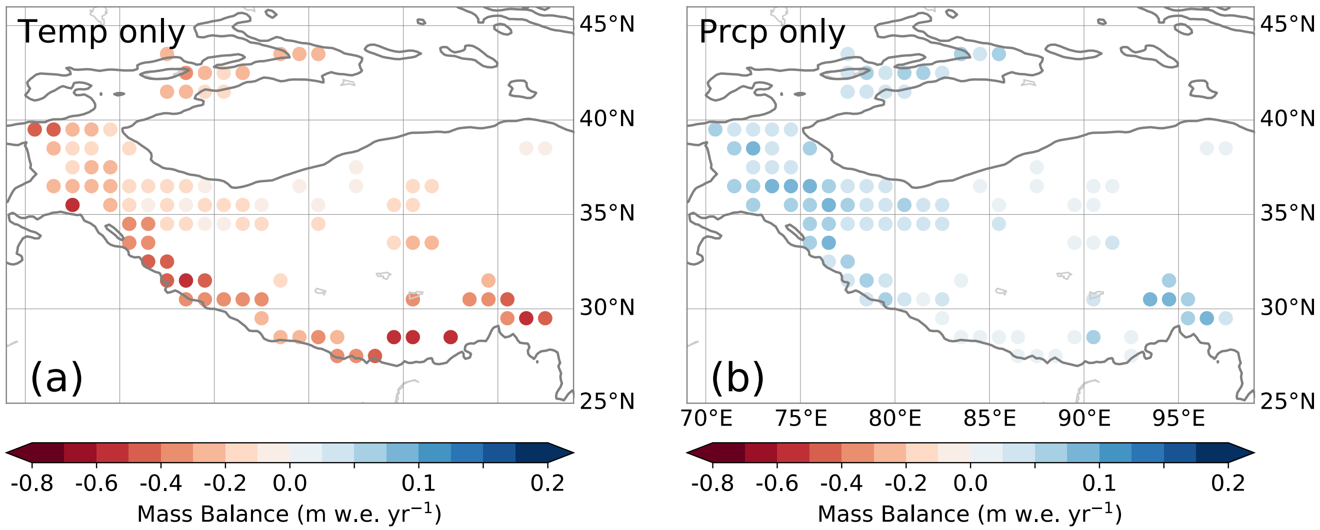


Figure 13: Simulated mean mass balance between 2000-2010 forced by a spatially uniform and constant temperature increase of +0.01 °C yr^{-1}, with snowfall kept constant (a), and a spatially uniform and constant snowfall increase of +0.5% yr^{-1} of the annual mean value, with temperature kept constant **(b). Panels *a* and *b* thus show the relative sensitivity to temperature and snowfall, respectively.**

3.4 Moisture sources

The trends in moisture source regions for WKSK (Fig. 14a,b) indicate that the largest increases in moisture from a given source to precipitation in WKSK occur in the mountains themselves. This increase in recycling occurs mainly in May, and is also the main cause of the increase in precipitation in September (see Fig. 15). The increase in recycling is probably a natural consequence of the increased precipitation there. The regions with the second largest increases are the areas in the Tarim basin where irrigation has increased the most, which contributes mainly in May-July, with May showing the largest resulting increase in snowfall (see Figs. 10 and 15). In July, the increase in Tarim irrigation still contributes to increasing precipitation in WKSK, but it falls more in the form of rain, compared to May, where it is mainly snow (Fig. 10). Another region that contributes to the increase in precipitation in WKSK is the Junggar basin, northeast of the Tarim basin. This is another arid region that has experienced rapid increases in irrigation. The increases per grid point are lower there, but they are spread out over a larger area. A final source region with an overall large positive trend is the Caspian Sea and the Caucasus. Note again that, due to a systematic offset in surface moisture flux between WRF and ERA-Interim, the moisture source trends in the Tarim and HMA are underestimated with respect to the other regions.

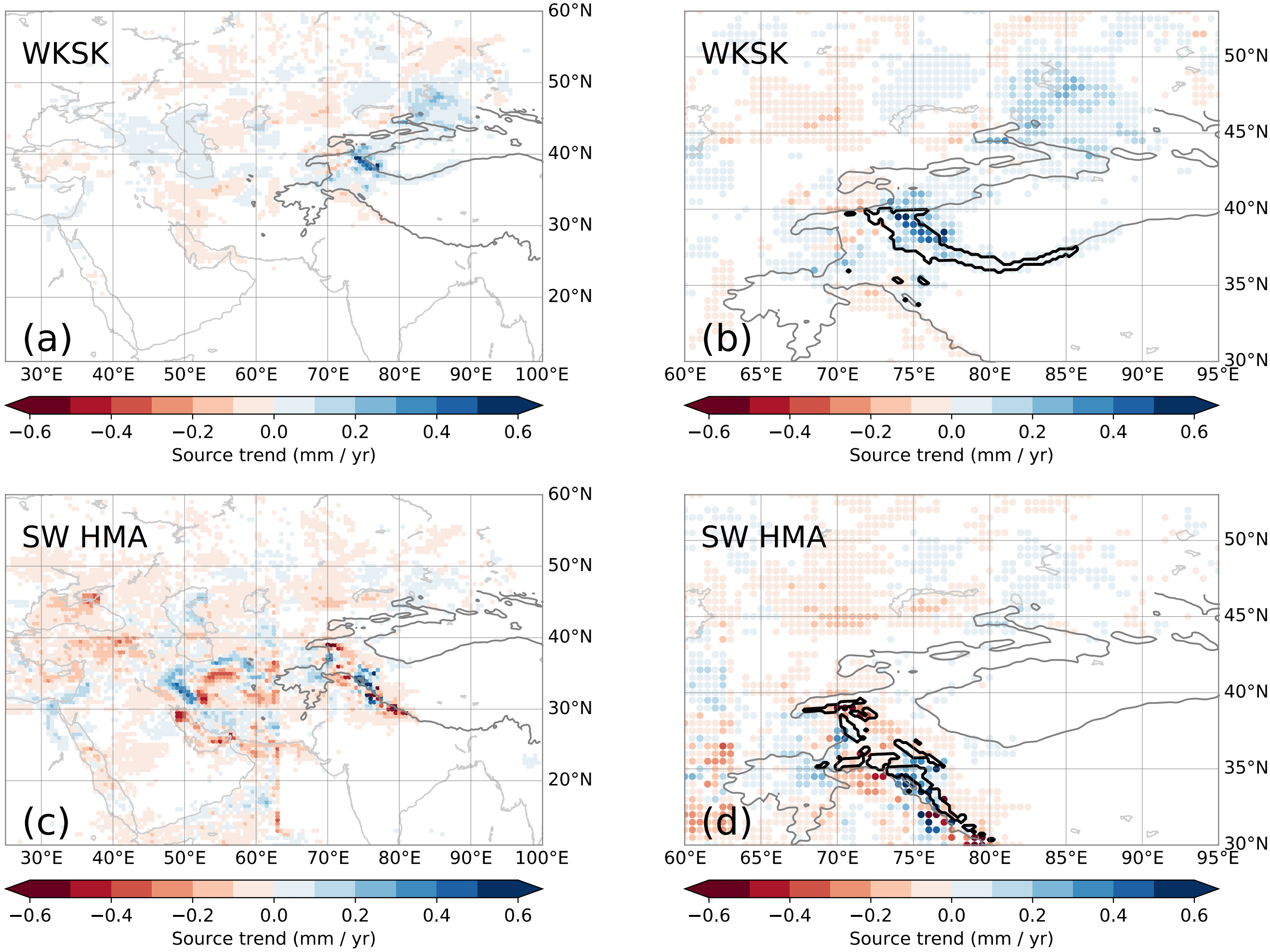


Figure 14: Trends in the amount of moisture from a given source contributing to precipitation trends in the target area (a,c), with a detailed view (b,d) around the target area from which the parcels were released (contoured in bold) for WKSK (a,b) and southwestern HMA (c,d). Trends with absolute magnitudes smaller than 0.02 mm yr^{-1} are made white. The 2000 m elevation contour in the WRF domain is indicated by a solid line.

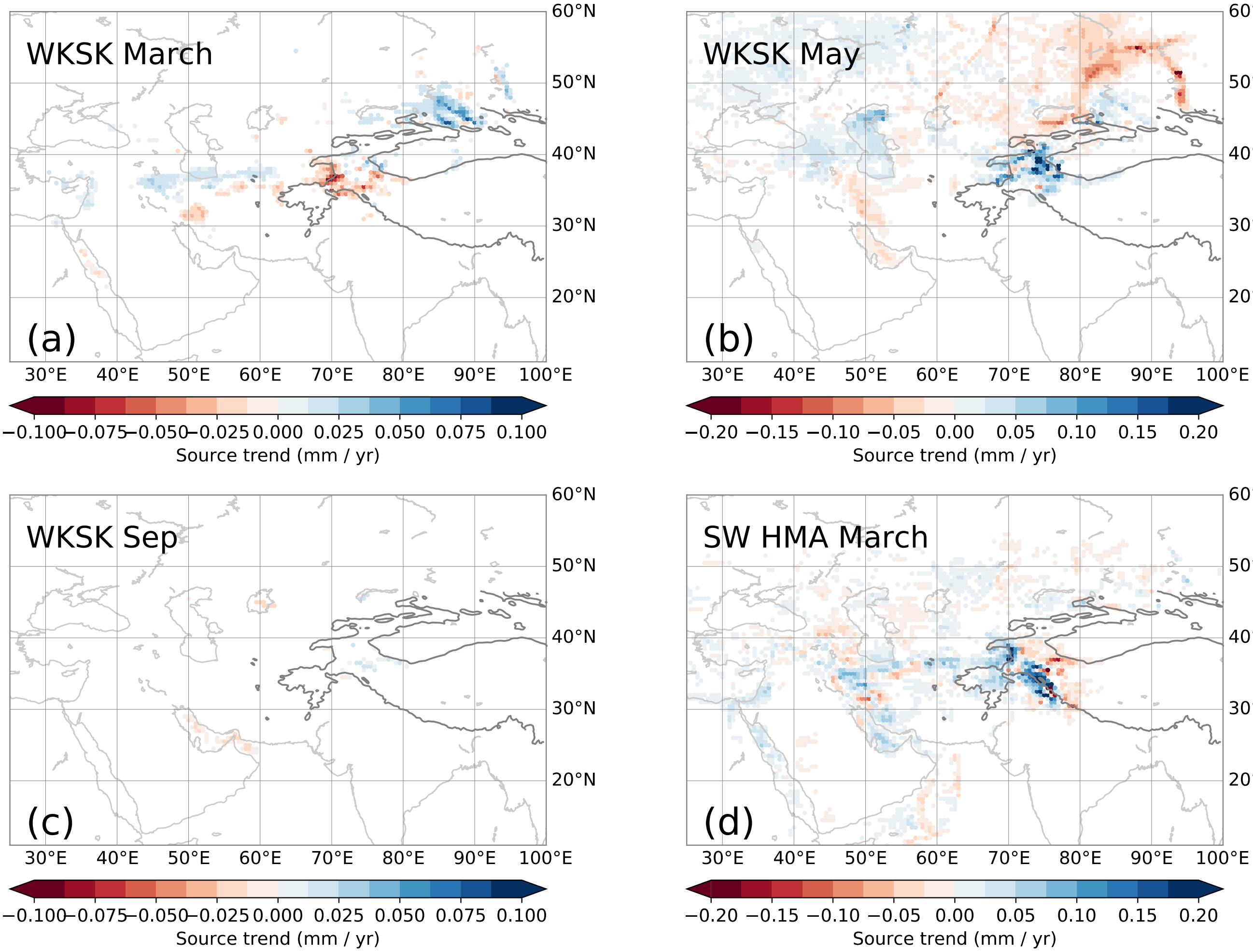


Figure 15: Trends in the amount of moisture from a given source for WKSK for March (a), May (b), September (c) and for southwestern HMA in March (d). These months correspond to large negative or positive trends in snowfall in Fig. 10. Note the different scales.

These results imply that evapotranspiration from irrigated areas in arid Northwest China play a large role in adding water to parts of HMA and hence to the observed positive mass balances. This is in line with recent work that shows that the recent wetting of Central Asia and the Tarim basin is associated with an increase in evapotranspiration in these regions (Dong et al., 2018; Peng et al., 2018; Peng and Zhou, 2017). The increase in the total evapotranspiration is influenced by the increase in potential evapotranspiration (Fang et al., 2018), increase in water availability (Jian et al., 2018), and increase in irrigated land area. On the interannual timescale, precipitation in WKSK strongly correlates with the moisture source amount in the western Tarim basin (Pearson r=0.96 below 3500 m, r=0.68 for the entire WKSK, as indicated in Fig. 14b). A similar correlation exists between the WKSK precipitation and the Caspian Sea moisture source amount (r=0.89 below 3500 m, r=0.43 for the entire

WKSK), showing the importance of the large-scale weather patterns. For the Junggar basin, this correlation is weaker (r=0.65 below 3500 m, r=-0.14 for the entire WKSK), since this region contributes relatively more in winter (Fig. 15), when less snowfall reaches WKSK (Fig. 10).

When performing the moisture tracking for the southwestern part of HMA, where snowfall has generally decreased (Fig. 14c,d), also the Caspian Sea and the Junggar basin positively contribute to the snowfall trend, whereas for these ranges the Tarim basin does not contribute to the snowfall trend, with maximal trends in moisture sources of less than 0.1 mm yr^{-1}. These results show that the irrigated areas in the Tarim basin are especially important in influencing the moisture supply to the Western Kunlun Shan (de Kok et al., 2018).

4. Conclusion and discussion

Our simulations, based on ERA-Interim and GLDAS reanalysis data, indicate that an increase of snowfall and a low temperature sensitivity are the main reasons why glaciers are growing or stable in western Kunlun Shan and Karakoram. This is the first time that the observed pattern of glacier mass balances in HMA is reproduced in a consistent way. **We show that such a pattern can be reproduced using relative changes in temperature and precipitation in recent decades. Since we used relative changes to force our glacier model, we are less influenced by errors in the absolute precipitation amounts, caused by our low resolution or by our choice of model physics. We illustrate this using WRF runs performed for de Kok et al. (2018) for May-September of two years. We ran WRF at two resolutions: at 20 km with same the physics settings as in this study, but without any nudging, and at 4 km, which is of high enough resolution to explicitly resolve convection and avoid the cumulus parameterisation. There are large local differences in precipitation between the two runs, mainly due to the difference in resolution. However, when the relative ratio of the precipitation is plotted for two years (Fig. 16), similar to what is used in the glacier model, the two set-ups give much more similar patterns. Snowfall gives very similar results, but we decided to show total precipitation, where total numbers and cumulus errors are expected to be even higher. The relative changes in precipitation do not markedly show the topography, in contrast to the individual precipitation fields. Rather, relatively large regions show similar interannual changes in the precipitation. The patterns of precipitation change also agree well between the 20 km results and the 4 km results, despite the very different treatment of the convection and the difference in topographic resolution. The differences between the scaling factor in the two cases can be of the order of tens of percent, which is much smaller than the difference in absolute precipitation amounts that would be needed to model the mass balance directly from the WRF fields. Also temperatures are mutually correlated over larger areas in WKSK (e.g. Forsythe et al., 2017) and the glacier mass balances in HMA also vary mainly over a large scale, suggesting that large-scale weather patterns are on average more important in controlling the interannual variability of temperature and precipitation than the differences between**

valleys. The use of relative changes in temperature and precipitation has thus made our results more robust against possible errors in the detailed treatment of the complex mountain meteorology.

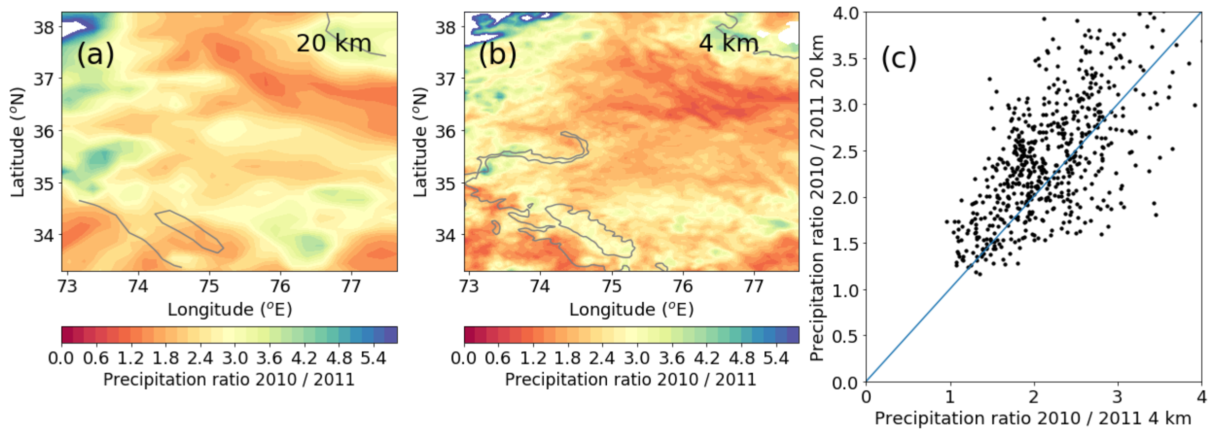


Figure 16: Precipitation ratios between May-September of two years for the WRF run at 20 km, with cumulus parameterisation (a), that at 4 km resolution, without parameterisation (b), and the two compared, when binned at the resolution of the 20 km run (c). The 2000 m elevation contour is indicated by a solid line.

One of our main sources of error is setting up the initial mass balance gradient, and our assumption that the glaciers are initially in balance. Due to the inertia of the glaciers, the initial condition has relatively large influence on the eventual mass balance decades later, as discussed above. Furthermore, any errors in the mass balance gradient, e.g. due to errors in the downscaling of ERA-Interim data, will affect the temperature and precipitation sensitivities presented here, but will have less impact on the overall pattern of mass balances in HMA, since they are mostly determined by the changes of temperature and precipitation.

Our snowfall trends between 1980-2010 show some similarities, but also major differences with respect to a similar WRF study that did not include irrigation and used another re-analysis dataset (Norris et al., 2018). For instance, our temperature trends do not exhibit the strong summer cooling at low altitudes (e.g. the Tarim basin), and are more in line with station data (Waqas and Athar, 2018; Xu et al., 2010) in that respect. However, contrasting precipitation trends in WKSK and southwestern HMA, similar to Fig. 8, are also present in ERA5 data and the Norris et al. study (see Farinotti et al., 2020). Although the interannual variability of temperature and precipitation is reasonably reproduced, and our precipitation trends are similar to those in other datasets, our model results are associated with uncertainties, which are partly irreconcilable due to a lack of *in situ* measurements in WKSK. Furthermore, different parameterisations in the regional climate model, different irrigation schemes, and different glacier models will likely yield slightly different results. Using an ensemble of such approaches could be used to assess the robustness of the results presented here in the future. Furthermore, detailed studies at smaller scales will give more insight into individual glacier behaviour. **It is then also possible to use more complex glacier models, e.g. those that take into account the full energy balance.**

The pattern of **snowfall** trends in Fig. 8b roughly matches the **precipitation** pattern that is expected from an increasing influence of summer westerlies, as shown by Mölg et al. (2017). **From this similarity, one could wonder whether the snowfall pattern from Fig. 8b is mainly caused by summer westerlies.** These **summer** westerlies are also associated with strong heating and drying trends of the Indus Basin. An increase in irrigation also produces a very similar precipitation pattern **as the pattern for summer westerlies**, yet causes a cooling and wetting of the Indus Basin (de Kok et al., 2018). Our JAS trends of near-surface temperature and specific humidity **from WRF** (Fig. 17) indicate mostly cooling and wetting trends **in the Indus basin**, which is more in line with the increase in irrigation than with the increase in summer westerlies. ERA5 data for JJA also indicates a similarly strong irrigation effect in the Indus basin (Farinotti et al., 2020), **as indicated by a wetting and cooling trend**. The moisture tracking results (Figs. 14 and 15) indicate that much of the additional snowfall occurs in spring and summer, and originates from the East, with a large role for the irrigated areas. The decrease in precipitation in southwestern HMA is also clearly associated with westerly winds in winter, but not those in summer (see Figs. 10d and 14c). The pattern of **snowfall** trends in Fig. 8b is **thus** not only the result of changes in summer. **When only JAS is considered, the pattern of precipitation trends look different from the annual snowfall trends (Fig. 17b). Therefore, the summer westerlies are likely not the main driver for the snowfall pattern seen in Fig. 8b.** However, the May westerlies clearly have an important role in transporting the increase in evaporation from the Caspian Sea (Chen et al., 2017) to WKSK. Besides the Caspian Sea, the westerlies are mainly associated with a decrease in snowfall when the whole year is considered (Fig. 14a).

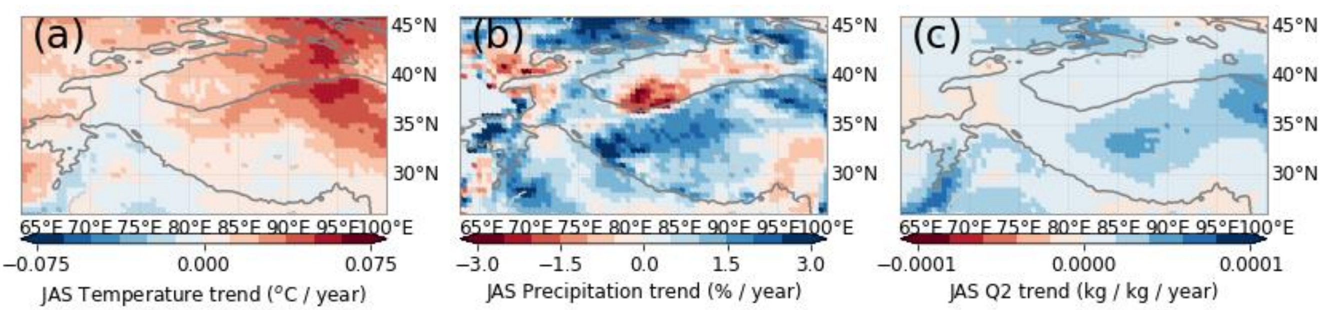


Figure 17: WRF trends between 1980-2010 of near-surface temperature (a), total precipitation (b), and specific humidity (b) between July-September, averaged over 0.5x0.5° bins for clarity. The 2000 m elevation contour is indicated by a solid line.

We show that the growing irrigated area in the arid region of Northwest China plays an important role in the increase in snowfall in WKSK. Previous studies have already shown that increases of irrigation in Northwest China can add precipitation to neighbouring mountains (Cai et al., 2019; de Kok et al., 2018), but we now show this process of increasing irrigation is also important compared to other changes in the atmosphere over the last few decades. Already before 1980, irrigation has increased in Northwest China (Fang et al., 2018), possibly contributing to the stable glacier conditions then. Future evolution of snowfall in this part of HMA is partly linked to how the irrigated areas develop in the future. Changes in temperature, irrigated area, or irrigation efficiency are therefore important parameters in understanding future run-off from glaciers and snow in WKSK. The increase in water availability for irrigation in Northwest China might be partly the result of the loss of glacier mass in Tien Shan (Dong et al., 2018). The mass loss will first result in an increase in glacier melt run-off into the Tarim basin, but ultimately the run-off will decrease as the glaciers shrink to a small size (Kraaijenbrink et al., 2017). On the other hand, if the primary source of irrigation water is groundwater, the amount of irrigation for the region will also have a limited sustainable or economic level. Once the groundwater is depleted, our results suggest that the glaciers in WKSK will also receive less snowfall from this region, resulting in their retreat. The relative importance of groundwater extraction, melt from Tien Shan, and recycling from WKSK, for water availability in the Tarim is yet unknown and will require future study. Furthermore, improving the estimates of irrigation gifts, e.g. by remote sensing, could also improve the past climate reconstruction of WKSK. Greening and warming in West-Asia could provide additional snowfall to WKSK, together with an increase in westerly disturbances (Cannon et al., 2015; Kapnick et al., 2014), but if temperatures in HMA keep increasing, the increase in melt will probably counteract glacier growth in most of HMA in the long term. **Our modelled mass balances show a decreasing mass balance trend for WKSK (Fig. 9d), but the trend is far too insignificant to draw conclusions about future mass balances.** It is clear that the coupling between glacier mass balance, runoff, and irrigation in different regions creates a complex problem of water availability, which will need to be researched further to inform decision makers on irrigation policies.

Acknowledgements

We acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant number 676819), the Netherlands Organisation for Scientific Research Innovational Research Incentives Schemes VIDI and VENI (016.181.308 and 016.171.019), and the Strategic Priority Research Program of Chinese Academy of Sciences (grant number. XDA20100300). Computing time was provided by the SURFsara CARTESIUS National Supercomputer of the Netherlands Organization for Scientific Research. We thank Rens van Beek for distribution of the PCR-GLOBWB irrigation data. We thank Fanny Brun and Jesse Norris for discussion. We thank Thomas Mölg, Dieter Scherer, and two anonymous reviewers for their careful reading and useful comments, which improved the manuscript.

Code and data availability

The data underlying our results in Figs. 8-15,17, i.e. monthly mean output from WRF of temperature and precipitation, annual glacier mass balances, and annual moisture sources, are directly accessible at *dataverse.nl (*https://hdl.handle.net/10411/ATONZD). Other data is available from the authors upon request. WRF and the glacier mass balance model are freely available. The moisture tracking model is available upon request from Obbe A. Tuinenburg.

Author contributions

R.J.d.K. and W.W.I. designed the study, with input from all authors. R.J.d.K. performed the WRF modelling, P.D.A.K. performed the glacier mass balance modelling, and O.A.T. performed the moisture tracking. All authors contributed to the writing and editing of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

References

AIRS Science Team and Teixeira, J.: AIRS/Aqua L3 Monthly Standard Physical Retrieval (AIRS-only) 1 degree x 1 degree V006, , doi:10.5067/Aqua/AIRS/DATA321, 2013.

Anon: C3S ERA5-Land reanalysis, Copernicus Clim. Chang. Serv. [online] Available from: https://cds.climate.copernicus.eu/cdsapp#!/home (Accessed 5 January 2020), 2019.

Archer, D.: Snow measurement, in Encyclopedia of Hydrology and Lakes. Encyclopedia of Earth Science., Springer, Dordrecht., 1998.

van Beek, L. P. H. and Bierkens, M. F. P.: The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification, Utr. Univ. Dep. Phys. Geogr. http//vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf [online] Available from: http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf, 2008.

van Beek, L. P. H., Wada, Y. and Bierkens, M. F. P.: Global monthly water stress: 1. Water balance and water availability, Water Resour. Res., 47(7), doi:10.1029/2010WR009791, 2011.

Beljaars, A. C. M.: The parametrization of surface fluxes in large???scale models under free convection, Q. J. R. Meteorol. Soc., 121(522), 255–270, doi:10.1002/qj.49712152203, 1995.

Bocchiola, D. and Diolaiuti, G.: Recent (1980-2009) evidence of climate change in the upper Karakoram, Pakistan, Theor. Appl. Climatol., 113(3–4), 611–641, doi:10.1007/s00704-012-0803-y, 2013.

Bolch, T., Kulkarni, a., Kaab, a., Huggel, C., Paul, F., Cogley, J. G., Frey, H., Kargel, J. S., Fujita, K., Scheel, M., Bajracharya, S. and Stoffel, M.: The State and Fate of Himalayan Glaciers, Science (80-.)., 336(6079), 310–314, doi:10.1126/science.1215828, 2012.

Bonekamp, P., de Kok, R., Collier, E. and Immerzeel, W.: Contrasting meteorological drivers of the glacier mass balance between the

Karakoram and central Himalaya, Front. Earth Sci., in press, doi:doi: 10.3389/feart.2019.00107, 2019.

Bonekamp, P. N. J., Collier, E. and Immerzeel, W.: The Impact of spatial resolution, land use, and spinup time on resolving spatial precipitation patterns in the Himalayas, J. Hydrometeorol., 19(10), 1565–1581, doi:10.1175/JHM-D-17-0212.1, 2018.

Brun, F., Berthier, E., Wagnon, P., Kääb, A. and Treichler, D.: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016, Nat. Geosci., 10(9), 668–673, doi:10.1038/ngeo2999, 2017.

Cai, P., Luo, G., He, H., Zhang, M. and Termonia, P.: Agriculture intensification increases summer precipitation in Tianshan, Atmos. Res., 227(April), 140–146, doi:10.1016/j.atmosres.2019.05.005, 2019.

Cannon, F., Carvalho, L. M. V., Jones, C. and Bookhagen, B.: Multi-annual variations in winter westerly disturbance activity affecting the Himalaya, Clim. Dyn., 44(1–2), 441–455, doi:10.1007/s00382-014-2248-8, 2015.

Chen, J. L., Pekker, T., Wilson, C. R., Tapley, B. D., Kostianoy, A. G., Cretaux, J. F. and Safarov, E. S.: Long-term Caspian Sea level change, Geophys. Res. Lett., 44(13), 6993–7001, doi:10.1002/2017GL073958, 2017.

Collier, E. and Immerzeel, W. W.: High-resolution modeling of atmospheric dynamics in the Nepalese Himalayas, J. Geophys. Res. Atmos., 120, 98822–9896, doi:10.1002/2015JD023266, 2015.

Cook, B. I., Shukla, S. P., Puma, M. J. and Nazarenko, L. S.: Irrigation as an historical climate forcing, Clim. Dyn., 44(5–6), 1715–1730, doi:10.1007/s00382-014-2204-7, 2015.

Copernicus Climate Change Service: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Clim. Chang. Serv. Clim. Data Store [online] Available from: https://cds.climate.copernicus.eu/cdsapp#!/home (Accessed 20 September 2005), 2017.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., H??lm, E. V., Isaksen, L., K??llberg, P., K??hler, M., Matricardi, M., Mcnally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Th??paut, J. N. and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137(656), 553–597, doi:10.1002/qj.828, 2011.

Dehecq, A., Gourmelen, N., Gardner, A. S., Brun, F., Goldberg, D., Nienow, P. W., Berthier, E., Vincent, C., Wagnon, P. and Trouvé, E.: Twenty-first century glacier slowdown driven by mass loss in High Mountain Asia, Nat. Geosci., 12, 22–27, doi:10.1038/s41561-018-0271-9, 2019.

Dlugokencky, E., Lang, P., Mund, J., Crotwell, M. and Thoning, K.: Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2017, , ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flas, 2018.

Dong, W., Lin, Y., Wright, J. S., Xie, Y., Ming, Y., Zhang, H., Chen, R., Chen, Y., Xu, F., Lin, N., Yu, C., Zhang, B., Jin, S., Yang, K., Li, Z., Guo, J., Wang, L. and Lin, G.: Regional disparities in warm season rainfall changes over arid eastern – central Asia, Sci. Rep., 8(August), 13051, doi:10.1038/s41598-018-31246-3, 2018.

Dyer, A. J. and Hicks, B. B.: Flux-gradient relationships in the constant flux layer, Q. J. R. Meteorol. Soc., 96(410), 715–721, doi:10.1002/qj.49709641012, 1970.

Van der Esch, S., Ten Brink, B., Stehfest, E., Bakkenes, M., Sewell, A., Bouwman, A., Meijer, J., Westhoek, H. and Van den Berg, M.: Exploring future changes in land use and land condition and the impacts on food, water, climate change and biodiversity Scenarios for the UNCCD Global Land Outlook Policy Report, PBL Rep. http//www.pbl.nl/sites/default/files/cms/publicaties/pbl-2017-exploring-future-changes-in-land-use-and-land-condition-2076.pdf [online] Available from: http://www.pbl.nl/sites/default/files/cms/publicaties/pbl-2017-exploring-future-changes-in-land-use-and-land-condition-2076.pdf, 2017.

Fang, G., Chen, Y. and Li, Z.: Variation in agricultural water demand and its attributions in the arid Tarim River Basin, J. Agric. Sci., 156(3), 301–311, doi:10.1017/S0021859618000357, 2018.

Farinotti, D., Immerzeel, W. W., Kok, R. J. De, Quincey, D. J. and Dehecq, A.: Manifestations and mechanisms of the Karakoram glacier

Anomaly, Nat. Geosci., 13(January), doi:10.1038/s41561-019-0513-5, 2020.

Forsythe, N., Fowler, H. J., Li, X.-F., Blenkinsop, S. and Pritchard, D.: Karakoram temperature and glacial melt driven by regional atmospheric circulation variability, Nat. Clim. Chang., (August), doi:10.1038/nclimate3361, 2017.

Fowler, H. J. and Archer, D. R.: Conflicting signals of climatic change in the upper Indus Basin, J. Clim., 19(17), 4276–4293, doi:10.1175/JCLI3860.1, 2006.

Gardelle, J., Berthier, E. and Arnaud, Y.: Slight mass gain of Karakoram glaciers in the early twenty-first century, Nat. Geosci., 5(5), 322–325, doi:10.1038/ngeo1450, 2012.

Gardelle, J., Berthier, E., Arnaud, Y. and Kääb, A.: Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999-2011, Cryosphere, 7(4), 1263–1286, doi:10.5194/tc-7-1263-2013, 2013.

Hewitt, K.: The Karakoram Anomaly? Glacier Expansion and the 'Elevation Effect,' Karakoram Himalaya, Mt. Res. Dev., 25(4), 332–340, doi:10.1659/0276-4741(2005)025[0332:TKAGEA]2.0.CO;2, 2005.

Hong, S.-Y., Noh, Y. and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes., Mon. Weather Rev., 134(9), 2318–2341, doi:10.1175/MWR3199.1, 2006.

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, J. Geophys. Res. Atmos., 113(13), 2–9, doi:10.1029/2008JD009944, 2008.

Immerzeel, W. W., Wanders, N., Lutz, A. F., Shea, J. M. and Bierkens, M. F. P.: Reconciling high-altitude precipitation in the upper Indus basin with glacier mass balances and runoff, Hydrol. Earth Syst. Sci., 19(11), 4673–4687, doi:10.5194/hess-19-4673-2015, 2015.

Jian, D., Li, X., Sun, H., Tao, H., Jiang, T., Su, B. and Hartmann, H.: Estimation of Actual Evapotranspiration by the Complementary Theory-Based Advection–Aridity Model in the Tarim River Basin, China, J. Hydrometeorol., 19(2), 289–303, doi:10.1175/JHM-D-16-0189.1, 2018.

Kääb, A., Treichler, D., Nuth, C. and Berthier, E.: Brief Communication: Contending estimates of 2003-2008 glacier mass balance over the Pamir-Karakoram-Himalaya, Cryosphere, 9(2), 557–564, doi:10.5194/tc-9-557-2015, 2015.

Kain, J. S.: The Kain–Fritsch Convective Parameterization: An Update, J. Appl. Meteorol., 43(1), 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2, 2004.

Kapnick, S. B. S., Delworth, T. L. T., Ashfaq, M., Malyshev, S. and Milly, P. C. D.: Snowfall less sensitive to warming in Karakoram than in Himalayas due to a unique seasonal cycle, Nat. Geosci., 7(11), 834–840, doi:10.1038/ngeo2269, 2014.

Khattak, M. S., Babel, M. S. and Sharif, M.: Hydro-meteorological trends in the upper Indus River basin in Pakistan, Clim. Res., 46(2), 103–119, doi:10.3354/cr00957, 2011.

de Kok, R. J., Tuinenburg, O. A., Bonekamp, P. N. J. and Immerzeel, W. W.: Irrigation as a Potential Driver for Anomalous Glacier Behavior in High Mountain Asia, Geophys. Res. Lett., 45(4), 2047–2054, doi:10.1002/2017GL076158, 2018.

Kraaijenbrink, P. D. A., Bierkens, M. F. P., Lutz, A. F. and Immerzeel, W. W.: Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers, Nature, 549(7671), 257–260, doi:10.1038/nature23878, 2017.

Lawrimore, J. H., Menne, M. J., Gleason, B. E., Williams, C. N., Wuertz, D. B., Vose, R. S. and Rennie, J.: Global Historical Climatology Network - Monthly (GHCN-M), Version 3, , doi:10.7289/V5X34VDR, 2011.

Lee, E., Sacks, W. J., Chase, T. N. and Foley, J. A.: Simulated impacts of irrigation on the atmospheric circulation over Asia, J. Geophys. Res. Atmos., 116(8), 1–13, doi:10.1029/2010JD014740, 2011.

Lin, H., Li, G., Cuo, L., Hooper, A. and Ye, Q.: A decreasing glacier mass balance gradient from the edge of the Upper Tarim Basin to the Karakoram during 2000 – 2014, Sci. Rep., (December 2016), 1–9, doi:10.1038/s41598-017-07133-8, 2017.

Lobell, D. B., Bonfils, C. and Faurès, J. M.: The role of irrigation expansion in past and future temperature trends, Earth Interact., 12(3), 1–11, doi:10.1175/2007EI241.1, 2008.

Lundquist, J., Hughes, M., Gutmann, E. and Kapnick, S.: Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks, Bull. Am. Meteorol. Soc., 100(12), 2473–2490, doi:10.1175/BAMS-D-19-0001.1, 2019.

Marshall, S. J., White, E. C., Demuth, M. N., Bolch, T., Wheate, R., Menounos, B., Beedle, M. J. and Shea, J. M.: Glacier water resources on the eastern slopes of the Canadian Rocky Mountains, Can. Water Resour. J., 36(2), 109–134, doi:10.4296/cwrj3602823, 2011.

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A. and Verhoest, N. E. C.: GLEAM v3: Satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10(5), 1903–1925, doi:10.5194/gmd-10-1903-2017, 2017.

Marzeion, B., Hock, R., Anderson, B., Bliss, A., Champollion, N., Fujita, K., Huss, M., Immerzeel, W., Kraaijenbrink, P., Malles, J., Maussion, F., Radić, V., Rounce, D. R., Sakai, A., Shannon, S., Wal, R. and Zekollari, H.: Partitioning the Uncertainty of Ensemble Projections of Global Glacier Mass Change, Earth's Futur., 8(7), e2019EF001470, doi:10.1029/2019ef001470, 2020.

Maurer, J. M., Schaefer, J. M., Rupper, S. and Corley, A.: Acceleration of ice loss across the Himalayas over the past 40 years, Sci. Adv., 5, 1–12, doi:10.1126/sciadv.aav7266, 2019.

Maussion, F., Scherer, D., Mölg, T., Collier, E., Curio, J. and Finkelnburg, R.: Precipitation seasonality and variability over the Tibetan Plateau as resolved by the high Asia reanalysis, J. Clim., 27(5), 1910–1927, doi:10.1175/JCLI-D-13-00282.1, 2014.

Ménégoz, M., Gallée, H. and Jacobi, H. W.: Precipitation and snow cover in the Himalaya : from reanalysis to regional climate simulations, Hydrol. Earth Syst. Sci., 17, 3921–3936, doi:10.5194/hess-17-3921-2013, 2013.

Miralles, D. G., Gash, J. H., Holmes, T. R. H., De Jeu, R. A. M. and Dolman, A. J.: Global canopy interception from satellite observations, J. Geophys. Res. Atmos., 115(16), 1–8, doi:10.1029/2009JD013530, 2010.

Mölg, T., Maussion, F., Collier, E., Chiang, J. C. H. and Scherer, D.: Prominent midlatitude circulation signature in high Asia's surface climate during monsoon, J. Geophys. Res. Atmos., 122(23), 12,702-12,712, doi:10.1002/2017JD027414, 2017.

Morrison, H., Thompson, G. and Tatarskii, V.: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes, Mon. Weather Rev., 137(3), 991–1007, doi:10.1175/2008MWR2556.1, 2009.

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M. and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, J. Geophys. Res. Atmos., 116(12), 1–19, doi:10.1029/2010JD015139, 2011.

Norris, J., Carvalho, L. M. V., Jones, C. and Cannon, F.: WRF simulations of two extreme snowfall events associated with contrasting extratropical cyclones over the western and central Himalaya, J. Geophys. Res. Atmos., 120(8), 3114–3138, doi:10.1002/2014JD022592, 2015.

Norris, J., Carvalho, L. M. V., Jones, C. and Cannon, F.: Deciphering the contrasting climatic trends between the central Himalaya and Karakoram with 36 years of WRF simulations, Clim. Dyn., 52(0), 159, doi:10.1007/s00382-018-4133-3, 2018.

östrem, G.: Ice Melting under a Thin Layer of Moraine, and the Existence of Ice Cores in Moraine Ridges, Geogr. Ann., 41(4), 228–230, doi:10.1080/20014422.1959.11907953, 1959.

Otte, T. L., Nolte, C. G., Otte, M. J. and Bowden, J. H.: Does nudging squelch the extremes in regional climate modeling?, J. Clim., 25(20), 7046–7066, doi:10.1175/JCLI-D-12-00048.1, 2012.

Palazzi, E., Von Hardenberg, J. and Provenzale, A.: Precipitation in the hindu-kush karakoram himalaya: Observations and future scenarios, J. Geophys. Res. Atmos., 118(1), 85–100, doi:10.1029/2012JD018697, 2013.

Paulson, C. a.: The Mathematical Representation of Wind Speed and Temperature Profiles in the Unstable Atmospheric Surface Layer, J.

Appl. Meteorol., 9(6), 857–861, doi:10.1175/1520-0450(1970)009<0857:TMROWS>2.0.CO;2, 1970.

Peng, D. and Zhou, T.: Why was the arid and semiarid Northwest China getting wetter in the recent decades?, J. Geophys. Res. Atmos., doi:10.1002/2016JD026424, 2017.

Peng, D., Zhou, T., Zhang, L. and Wu, B.: Human Contribution to the Increasing Summer Precipitation in Central Asia from 1961 to 2013, Clim. Dyn., 31, 8005–8021, doi:10.1175/JCLI-D-17-0843.1, 2018.

Pfeffer, W. T., Arendt, A. A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A. S., Hagen, J. O., Hock, R., Kaser, G., Kienholz, C., Miles, E. S., Moholdt, G., M??lg, N., Paul, F., Radi??, V., Rastner, P., Raup, B. H., Rich, J., Sharp, M. J., Andreassen, L. M., Bajracharya, S., Barrand, N. E., Beedle, M. J., Berthier, E., Bhambri, R., Brown, I., Burgess, D. O., Burgess, E. W., Cawkwell, F., Chinn, T., Copland, L., Cullen, N. J., Davies, B., De Angelis, H., Fountain, A. G., Frey, H., Giffen, B. A., Glasser, N. F., Gurney, S. D., Hagg, W., Hall, D. K., Haritashya, U. K., Hartmann, G., Herreid, S., Howat, I., Jiskoot, H., Khromova, T. E., Klein, A., Kohler, J., K??nig, M., Kriegel, D., Kutuzov, S., Lavrentiev, I., Le Bris, R., Li, X., Manley, W. F., Mayer, C., Menounos, B., Mercer, A., Mool, P., Negrete, A., Nosenko, G., Nuth, C., Osmonov, A., Pettersson, R., Racoviteanu, A., Ranzi, R., Sarikaya, M. A., Schneider, C., Sigurdsson, O., Sirguey, P., Stokes, C. R., Wheate, R., Wolken, G. J., Wu, L. Z. and Wyatt, F. R.: The randolph glacier inventory: A globally complete inventory of glaciers, J. Glaciol., 60(221), 537–552, doi:10.3189/2014JoG13J176, 2014.

Prinn, R. G., Weiss, R. F., Simmonds, P. J. F. P. G., Cunnold, D. M., Alyea, F. N., Doherty, S. O., Salameh, P., Miller, B. R., Huang, J., Wang, R. H. J., Hartley, D. E., Harth, C., Steele, L. P., Sturrock, G., Midgley, P. M. and Mcculloch, A.: A history of chemically and radiatively important gases in air deduced from ALE_GAGE_AGAGE, J. Geophys. Res. Atmos., 105(D14), 17,751-17,792, 2000.

Puma, M. J. and Cook, B. I.: Effects of irrigation on global climate during the 20th century, J. Geophys. Res. Atmos., 115(16), 1–15, doi:10.1029/2010JD014122, 2010.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin*, J. K., Walker, J. P., Lohmann, D. and Toll, D.: The Global Land Data Assimilation System, Bull. Am. Meteorol. Soc., 85(3), 381–394, doi:10.1175/BAMS-85-3-381, 2004.

Sacks, W. J., Cook, B. I., Buenning, N., Levis, S. and Helkowski, J. H.: Effects of global irrigation on the near-surface climate, Clim. Dyn., 33(2–3), 159–175, doi:10.1007/s00382-008-0445-z, 2009.

Sakai, A. and Fujita, K.: Contrasting glacier responses to recent climate change in high-mountain Asia, Sci. Rep., 7(1), 13717, doi:10.1038/s41598-017-14256-5, 2017.

Senay, G. B.: Satellite psychrometric formulation of the operational simplified surface energy balance (SSEBOP) model for quantiying and mapping evapotranspiration, Appl. Eng. Agric., 34(3), 555–566, doi:doi: 10.13031/aea.12614, 2018.

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P. and Portmann, F. T.: Groundwater use for irrigation - A global inventory, Hydrol. Earth Syst. Sci., 14(10), 1863–1880, doi:10.5194/hess-14-1863-2010, 2010.

Skamarock, W. C. and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, J. Comput. Phys., 227(7), 3465–3485, doi:10.1016/j.jcp.2007.01.037, 2008.

Susskind, J., Blaisdell, J. M. and Iredell, L.: Improved methodology for surface and atmospheric soundings, error estimates, and quality control procedures: the atmospheric infrared sounder science team version-6 retrieval algorithm, J. Appl. Remote Sens., 8(1), 084994, doi:10.1117/1.jrs.8.084994, 2014.

Tuinenburg, O. A., Hutjes, R. W. A. and Kabat, P.: The fate of evaporated water from the Ganges basin, J. Geophys. Res. Atmos., 117(1), 1–17, doi:10.1029/2011JD016221, 2012.

Ul Hasson, S., Böhner, J. and Lucarini, V.: Prevailing climatic trends and runoff response from Hindukush-Karakoram-Himalaya, upper Indus Basin, Earth Syst. Dyn., 8(2), 337–355, doi:10.5194/esd-8-337-2017, 2017.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald,

A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A. Pietro, Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C. N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D. A., Hagen, D. R., Pasechnik, D. V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G. A., Ingold, G. L., Allen, G. E., Lee, G. R., Audren, H., Probst, I., Dietrich, J. P., Silterra, J., Webber, J. T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J. L., de Miranda Cardoso, J. V., Reimer, J., Harrington, J., Rodríguez, J. L. C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N. J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P. A., Lee, P., McGibbon, R. T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., et al.: SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods, 17(3), 261–272, doi:10.1038/s41592-019-0686-2, 2020.

Wada, Y., Van Beek, L. P. H., Viviroli, D., Drr, H. H., Weingartner, R. and Bierkens, M. F. P.: Global monthly water stress: 2. Water demand and severity of water stress, Water Resour. Res., 47(7), 1–17, doi:10.1029/2010WR009792, 2011.

Wada, Y., Wisser, D. and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth Syst. Dyn., 5(1), 15–40, doi:10.5194/esd-5-15-2014, 2014.

Wang, R., Liu, S., Shangguan, D., Radic, V. and Zhang, Y.: Spatial Heterogeneity in Glacier Mass-Balance Sensitivity across High Mountain Asia, Water, 11, 776, doi:10.3390/w11040776, 2019.

Waqas, A. and Athar, H.: Recent decadal variability of daily observed temperatures in Hindukush , Karakoram and Himalaya region in northern Pakistan, Clim. Dyn., 0(0), 0, doi:10.1007/s00382-018-4557-9, 2018.

Webb, E. K.: Profile relationships: The log-linear range, and extension to strong stability, Q. J. R. Meteorol. Soc., 96(407), 67–90, doi:10.1002/qj.49709640708, 1970.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J. and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, Water Resour. Res., 50, 7505–7514, doi:10.1002/2014WR015638, 2014.

Xu, Z., Liu, Z., Fu, G. and Chen, Y.: Trends of major hydroclimatic variables in the Tarim River basin during the past 50 years, J. Arid Environ., 74(2), 256–267, doi:10.1016/j.jaridenv.2009.08.014, 2010.

Zhang, D. and Anthes, R. A.: A High-Resolution Model of the Planetary Boundary Layer—Sensitivity Tests and Comparisons with SESAME-79 Data, J. Appl. Meteorol., 21(11), 1594–1609, doi:10.1175/1520-0450(1982)021<1594:AHRMOT>2.0.CO;2, 1982.