General comments

The paper from Shaw et al. presents a new interesting dataset of on-glacier air temperature measurements, which are compared to local off-glacier measurements in order to estimate the climatic sensitivity along the glaciers, and its spatial and temporal variability. In addition, the authors compare their results with similar datasets collected around the world, making assessments on the generalizability of the Shea and Moore (2010) approach.

The new dataset is of relevance, also because it provides additional observations for long flowline distances, which have been undersampled so far. The results are interesting and substantially agree with former parameterizations, confirming their good generalizability. This is a nice intercomparison among different sites in different geographic areas, even if it is not the first as suggested the Introduction.

In my opinion the paper is rather well written, even if there are parts that require rewriting to be clear. I refer the authors to the specific comments, but I also suggest a careful proof reading (sorry I am not a native speaker).

One of my major concerns is related to the 'accuracy' assessment of the air temperature measurements. I have written a specific comment on this regard. In general, the authors compare measurement datasets collected in the last decades using rather different experimental setups, in particular for radiation shields that range from aspirated research-grade ones, to simple passive-ventilation tubes opened at both ends. The authors are well aware of this issue, as they write in the discussion and the outlook. I am not criticizing the single experimental designs and approaches, but I agree with them that this is one of the main issue for intercomparisons, often overlooked, and that in the future this aspect requires higher attention.

According to me, the authors should add some discussions on the influence of local glacier surface topography, and in particular the width/length ratio of glacier tongues (i.e. the heating from surrounding slopes that are free from ice cover) and the role of surface steepness upslope of Ta sites. Clearly, the Shea and Moore (2010) does not account for glacier slope, which however is important in regulating the local prevalence of the cooling caused by the loss of sensible heat over the adiabatic heating for descending air. It could be interesting to observe that the Greuell and Böhm (1998) approach accounts for this.

Specific comments

Title: please consider this or similar alternatives (less generic): Distributed summer air temperatures across three glaciers in the Tibetan Plateau: climatic sensitivity and comparison with existing glacier datasets

30: both the Ta 'on' (or over) the Tibetan glaciers

31: In general, observations on small glaciers (*with flowline shorter than 1000 m*) *reveal that they* are highly sensitive to temperature changes outside the glacier boundary layer. Or maybe I have misunderstood the meaning, a possible alternative could be: In general, *air temperature* on small glaciers (*with flowline shorter than 1000 m*) *is* highly sensitive to temperature changes outside the glacier boundary layer.

35: please rephrase, e.g. replace 'remain associated with other warm air' with 'i*s affected by warming*' processess that increase…

40: is still required to explain '*the variability'* of these effects '*on*' different glaciers

48-49: that or which?

51: maybe better 'increases' rather than 'heightens'?

74-76: please rephrase to improve readability of this sentence (and in general of the paper, adding some commas, or splitting long sentences)

77-78: have been tested by …..

82-86: this period requires commas and/or splitting

88: adds, additional, is repetition ('additional' should be removed in my opinion, also because there is another one at the end of the period)

95: consider replacing 'Thus' with 'In this way'. The separation from the period above is questionable in my opinion

98: Because distributed on-glacier observations are often limited, and rarely cover the full length of the glacier boundary layer (or flowline?), this additional correction for warming associated to the unknown parameters (am I right?) of ModGB can lead to high uncertainty in Ta estimates at the glacier terminus (or maybe better 'lower part'?)

104: to be the cause of 'the' relative warming 'at' the glacier terminus

107-109: in my opinion, the simplicity and the statistical nature of an approach are not sufficient by themselves to make it more generalizable than more complex schemes. Maybe you meant 'easily applicable'? The generalizability of the Shea and Moore (2010) approach was instead assessed by previous works and applications to other glaciers. Please consider revising.

119: on-glacier observations available around the world. (I would delete 'made to date').

111-112: I do not agree with this statement. Please see for example the comparisons made by Greuell and Böhm (1998) in their Figure 5, and by Carturan et al., (2015) in their figures 6 and 8

130: 'wider Parlung catchment' is not so clear. Please see also my comments in this regard for Figure 2.

137: and the 'longest' rain season? (I would remove 'annual')

141: please clarify what is the 'more elevation-independent mass balance sensitivity' (lower sensitivity of mass balance to elevation?)

142: 'Because Tibetan glaciers are shrinking and fragmenting, the accurate estimation of on-glacier temperatures is relevant for investigating and modelling their climatic sensitivity (Carturan et al., 2015). However, to date, no studies regarding the distribution of on-glacier temperature have been performed within the Tibetan Plateau'

149: in this section you report the manufacturer and model of temperature sensors but information on manufacturer and model of radiation shields is of equal (or even larger) importance. Please add this information.

156: from 12th July 'to' 18th September (exactly the same day/month in both years?)

157: here you mention Table 1 after speaking about data gaps, but Table 1 does not report this information, please check or modify

163-164: the sentence from 'For' to 'station' leaves the reader with incomplete information on the spatialization of off-glacier air temperature. Is it in the following sections? If so, move this sentence there; if not, complete here

169: in my opinion this intercomparison is useful, but it does not reveal the 'uncertainty' of measurements, which would require an intercomparison among instruments and shields under controlled conditions (see

e.g.: https://www.wmo.int/pages/prog/www/IMOP/PastIntercomparisons.html). Instead, it enables to assess the 'comparability' of measurements, which is relevant in this work but is different from uncertainty (it would require at least a comparison towards an artificially-ventilated radiation shield). I want to highlight this difference because both $T4_4$ and AWS_On have passive-ventilation systems, which are subject to the same error under similar conditions. For example, they are likely to lead to significant overestimation of air temperature under high radiation and low wind speed. In these circumstances a small difference between the two instruments does not reveal that they are affected by small uncertainty, it only enables to state that they are comparable. I have measured differences up to several degrees in these conditions between passive- and active-ventilation systems on glaciers, especially with high surface albedo. This is the reason why I prefer using active-ventilation systems, even if I perfectly understand the advantage of using passive ventilation systems, which in harsh high-mountain environments are often the only option (also considering costs, power requirement and possibility of failure of active systems).

170: in table 1, $T4_4$ is at 4809 m (same elevation of $T3_4$), whereas AWS_On is at 4649 m. On the other hand, in Figure 2$T4_4$ and AWS_On are at the same place, therefore I think there is an error in Table 1

171: that are located (remove 'co-')

174: this 'P90' is not clear, it looks like an alternative name for the AWS_Off. Please clarify, for example you could write: '……temperature at AWS_Off. We find that for these warm hours (hereafter referred to as 'P90' - (Ayala et al., 2015; Shaw et al., 2017; Troxler et al., 2020)), when the KBL development is theoretically at its strongest (e.g. van den Broeke, 1997; Oerlemans and Grisogono, 2002), 95% of hourly differences……'

183: based on these results, did you consider filtering the temperature data, e.g. excluding those with ERA5 wind speed lower that 2 m/s?

186-188: 'In addition to Ta, in this work we used incoming shortwave radiation and relative humidity measured at AWS_Off, on-glacier wind speed measured at AWS_On, and 'free-air' wind speed and direction derived from ERA5 reanalysis data (C3S, 2017)'.

192: to obtain elevation information

201: please remove 'that' before 'the impact'

202: at few 'selected' points

206: 'for aggregation of on-glacier ta data into sub-groups, and for calculating the distribution of off-glacier ta in space.

209: and 'a quantification of' the effect of the glacier bl. (is referred to the 'allows' above)

212: are 'described' in sections…

214: please state explicitly to what P10 and P90 correspond, i.e. the 10% coldest and the 10% warmest off-glacier temperature? and what about data between them?

217-219: unclear

219-222: can be written more clearly, e.g. add 'on-glacier' to Ta (be explicit)

234: consider removing 'bias'

278-279: I do not understand this sentence from 'we' to 'site' (maybe a repetition of what is reported at L286-287?)

299: what is ‘the best available value from the relevant literature’ for the annual precipitation at the AVDM site? Carturan et al., (2015) report 1233 mean annual precipitation at that site from 1979 to 2009, therefore I wonder where the authors found 784 mm as reported in Table 2 (it looks like uncorrected raw data from the Careser diga weather station….). In 2010 and 2011 the annual precipitation was above average (i.e. about 1400 mm).

308: please report n for all hours

309: flowline ‘distance’ (or DF), also in the following

312: and ‘with’ flowline distance

321: when comparing ‘Ta’ to glacier elevation or flowline (distance)?

324: please add ‘at the same elevation’ after ‘Ta’

335: ‘i.e. where on-glacier observations are expected to match Taamb’

340-342: I understand the meaning but this sentence should be written more clearly, e.g. there is a good alignment between on/off-glacier temperature until the katabatic winds start to blow on Parlung4

345: and ‘extension’ of the katabatic wind into the proglacial area

347: also here consider removing ‘bias’ (in any case I would use bias or offset, but not both). In the following you use only offset so I would stick to that

351: These effects appear to ‘increase’ beyond 2000 m 352 along the flowline on the Parlung94, though ‘significant’ offsets are observed for all glaciers (and sites?)

360: would it be possible to state that, besides individual meteorological variables, the offset is largest for warm/anticyclonic conditions, and lowest for cool/cyclonic conditions?

364: the terminus ‘of’ each glacier

366: recorded at AWS_off

369: though ‘it’ varies….

371: ‘…September, while the offset for Parlung4 and Parlung94 remain significant’. Possible explanation for that?

376-387: this part is written rather poorly, please improve it (concepts are more clear and better described e.g. in the caption of Figure 8). In particular, try to avoid writing ‘parameterisations’ (too general) and write explicitly whether you are referring to ‘sensitivities’ or ‘fitted exponential functions’

388-396: these results are interesting, and I wonder if there can be some influence from the very low width/length ratio of the tongues of McCall and Juncal Norte, surrounded by ice-free slopes that are expected to warm considerably… Parlung4 instead has a more ‘compact’ shape (i.e. less thin and long tongue). In addition I ask the authors to consider possible differences in the sensors and radiation shields used in the various referenced works and datasets (e.g. the radiation shields used in Juncal Norte were PVC cylinders opened at both ends, that are very different from the generally-used multi-plate radiation shieds)

399: the investigated sites lie close to the original SM10 exponential function up to ~4000 m….

416-423: poorly written, please clarify (requires multiple reading)

427: bring additional evidence of the spatial…

429: and highlights the need to appropriately account for these effects in glacier melt models

431: remove ‘since’?

440: To the authors’ knowledge?

487: remove ‘more’ before ‘generally following…’

489: ‘in agreement with the ‘ModGB’ model…’

499: the most distant station: the station with the largest flowline distance?

564: for a more limited number of total observations: do you mean ‘from a smaller sample size of temperature observations’?

566: however we found negligible sensitivity of derived K2 on T* (If I understand the meaning correctly)

575: rise or onset or increasing strength? I would rewrite, e.g.: However, ignoring the differences in climatic sensitivity that have been reported before and after the onset of the KBL, is arguably an over-simplification and does not enable to describe correctly the observed behaviour

578-584: poorly written

588: please remove ‘more’ after exploring. Carturan et al. (2015) used these ventilated devices, are there other studies with ventilated shields among those compared in this work? I would add here the related references

591: ‘at’ the glacier terminus

593: for certain glaciers or glacier areas?

606: please replace ‘are more climate sensitive’ with ‘display high climatic sensitivity’

607: that is close to 1 (remove ‘r’ at the end of closer, otherwise you should report the term of comparison)

618: compared to the existing parameterization

623: is associated with (not ‘remain’)

624: I would add (also in the discussion) the thermal (infrared) emission from the surrounding bare rocks, especially for thin and long ablation tongues

627: this is not true if the tongue is thin (small % area compared to the total glacier area). I would rewrite: is highly important because this is the glacier part that is most affected by ablation

633: to enable the construction of a generalised model for on-glacier air temperature estimation.

Figure 1: in my opinion the green parts of fig. 1a can be removed. I would also change the colour of (a), (b) and (c) from red to black, and try to uniform their placement (e.g. upper right corner of their frames). I would also shorten (and make more easily readable) the caption. The difference between (b) and (c) is not well explained. In a general scheme like this, I think that one should also find a short explanation of what k1, k2 and the yellow star represent

Figure 2: in b) I cannot see the map of a catchment, i.e. one expects to see the boundary of a hydrological catchment, with its divide, rivers and outlet. Please rephrase. When you mention suffix please refer to table 1. Please note that (A) does not show elevation data (it is (b)). c-e show the calculated flowline distances. Please consider homogenizing the scales for flowline distances. Off-glacier T2 on Parlung4 and T3 on Parlung94 actually look like ‘on-glacier’ in the figure.

Figure 3: caption ‘The mean Ta against elevation…’ (see caption Figure 4)

Figure 4: title of second chart (P90) is partially cut. Caption: ‘The red line indicates the piecewise lapse rate above the elevation of T1_390 ~~to lapse Ta~~ to the top of the flowline’

Figure 5: please write explicitly ‘wind speed’ in the y axis (WS is the acronym but generates confusion with the WS of ‘AWS’ to a quick reader). Caption: you can remove the sentence from ‘No’ to ‘data’, it is sufficient to describe it in the maintext.

Figure 6: why are the legends semi-transparent? Caption: (estimated ‘from Taamb’ - observed). high ‘is’ > 2.5 m s-1 and low ‘is’ < 0.7 m s-1)

Figure 7: in the caption: Maximum daily Ta offsets (estimated ‘from Taamb’ - observed) at the on-glacier T-Logger closest to the terminus of each glacier for 2018 (missing!) (top) and 2019 (bottom).

Figure 9: Caption: The original SM10 parameterisation is retained in the top panels

Figure 10: The k2 sensitivity along the normalized flowline compared to the total glacier length (colour bar). Glaciers have been grouped in two clusters: a) those with downglacier decreasing sensitivity, and b) those with increasing sensitivity towards the glacier terminus.

Figures 9 and 10: I find difficult to refer to Figure 8 for site coding (consider repeating the legend with geometric symbols in these figures)

Table 2: please check the precipitation data (e.g. AVDM and the 12000 mm in Chile). At line 981: ‘Annual’ precipitation totals….. Moreover, to which elevation are referred the MSAT and PT? I think it is relevant considering the high vertical lapse rates for both variables.

984: and those ‘fitted’