Supplement to: Multi-physics ensemble snow modelling in the western Himalaya

S1 Climate Input Strategies

The baseline climate inputs for this study are based primarily on downscaling and bias correction (for temperature) of the High Asia Refined Analysis (HAR, Maussion et al., 2014) dynamical downscaling product, as described in Section 3.2.
5 Given the uncertainties in climate input fields in this data-sparse context, simulations were also performed using two alternative input derivation strategies (Section 3.2.3). These strategies are summarised in Table S1 below. The strategies are not independent, as their main purpose is to indicate whether the conclusions reached on snowpack process representations, the focus of this study, are unduly affected by the downscaling and bias correction approaches described in Section 3.2.2. Precipitation is kept consistent between strategies, as the HAR represents by far the best available source of distributed
10 precipitation fields (Pritchard et al., 2019). The focus is thus on climate variables used in surface energy balance calculations. The implications of using these alternative input strategies are discussed in Section 5 and Section S5.

Input Strategy	Description
1	Downscaled HAR using approach described in Section 3.2.2, with simple bias correction of temperature fields. This is considered to be the best approach with available data and thus the baseline for the study.
2	Downscaled HAR using approach described in Section 3.2.2, but without any bias correction of temperature fields. The primary purpose of this strategy (2) is to check whether the temperature bias correction applied in (1) alters inter-variable relationships in a way which affects ensemble structure.
3	Downscaled HAR precipitation as per Section 3.2.2, but with other climate fields estimated primarily from observations. Specifically, temperature is lapsed based on observations (separately for daily minima and maxima, using monthly lapse rates). Daily temperatures are disaggregated to an hourly interval based on normalised climatological hourly diurnal cycle from EvK2CNR stations for each month. Relative humidity is estimated from daily minimum and maximum observations, and disaggregated to an hourly time step using a similar approach. Incoming shortwave radiation is calculated as per strategies (1) and (2). However, rather than estimating cloud transmissivity from the HAR (Section 3.2.2), a parameterisation of cloud transmissivity based on diurnal temperature range is used following calibration with local data (Pellicciotti et al., 2011). Incoming longwave radiation is estimated using the formulation from MicroMet (Liston and Elder, 2006). Wind speed is taken from the HAR in the absence of observations.

Table S1. Summary of baseline and alternative climate input data sources and strategies.

15 **S2 MODIS Land Surface Temperature (LST) Validation**

Figure S2 compares MODIS MOD11A1 Collection 6 land surface temperature (LST) with observations from the EvK2CNR Concordia site (Figure 1b), in order to provide further validation of the remote sensing product in this region (Section 3.3). Observed LST was estimated from measured outgoing longwave radiation at Concordia for the hours closest to the MODIS overpass times. The corresponding MODIS LST values were based on the average of the 9 pixels surrounding
20 a station location and were corrected for elevation differences using local MODIS LST lapse rates (estimated from linear regression). Figure S2 shows that the MODIS LST shows good correspondence with observations overall, as reflected by the values lying generally close to the 1:1 line. The summary statistics in Table S1 confirm that the MODIS bias is low at the annual scale (-1.6°C for night-time and 0.5°C for daytime), although it may be slightly larger for individual seasons. Nevertheless, Table S1 indicates that MOD11A1 is likely accurate enough to estimate climatological LST to within 2-3°C,
25 depending on season.
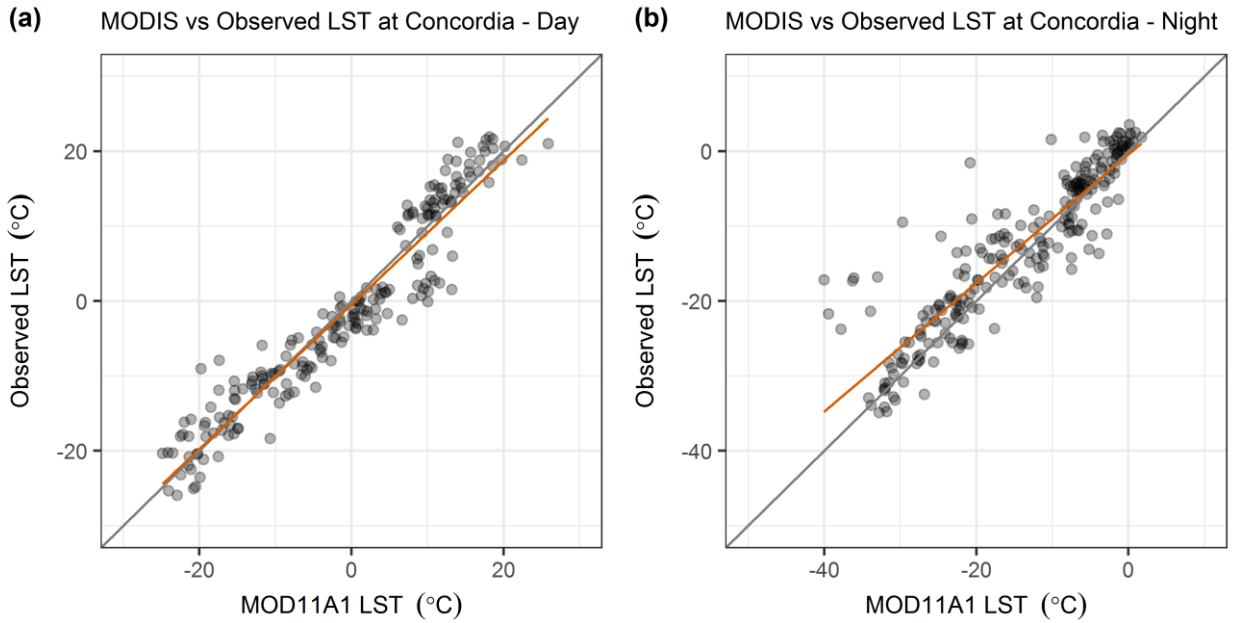
Figure S1. Comparison of (a) daytime and (b) night-time LST from MODIS remote sensing (MOD11A1) with observations at the Concordia site. Observed LST was derived from measured outgoing longwave radiation for the hour closest to the MODIS overpass time. The orange lines represent best fits from linear regression.

Season	Bias (°C)		RMSE (°C)	
	Night	Day	Night	Day
Annual	-1.6	0.5	5.0	3.7
DJF	-3.5	-1.4	3.4	7.2
MAM	-0.2	2.1	3.6	4.7
JJA	-0.7	1.4	5.0	3.7
SON	-2.8	-1.4	2.2	4.3

Table S2. Summary statistics for MOD11A1 performance at the EvK2CNR Concordia site.

S3 Model Performance Trade-Offs

Figure S2 shows a scatterplot matrix comparing root-mean-square deviation (RMSD) for ensemble groups using
10 different configurations of albedo (A), liquid water / drainage (D) and stability adjustment (S) options. This expands the number of groups shown in Section 4.3.1. The upper right of the scatterplot matrix shows runoff RMSD, whereas the lower left shows SCA RMSD. Each point represents the RMSD for a single year, and results from all three climate input strategies are plotted (Section 3.2.3 and Section S1). Figure S2 highlights the variation of performance relationships between ensemble groups. For both runoff and SCA, the most linear relationships tend to be associated with configurations using diagnostic
15 albedo (A0). This suggests that this choice dominates performance when selected. In some other cases, possible Pareto-like fronts are accompanied by almost linear relationships away from the front (examples for runoff include A1D1S0 vs A1D0S0 and A1D0S1). This indicates that the relative importance of differences between ensemble groups may vary between years in some cases.
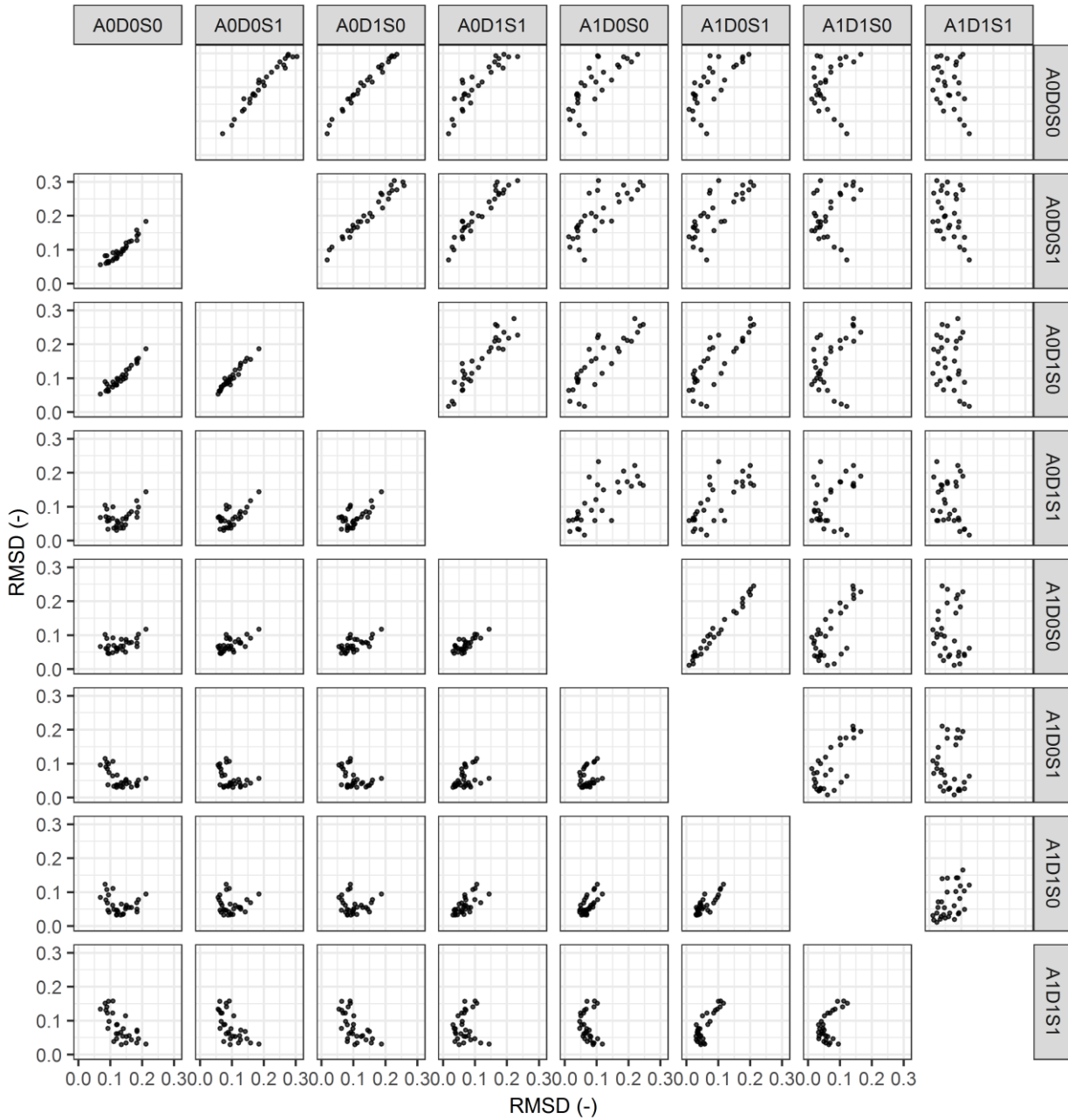
Figure S2. Scatterplot matrix comparing root-mean-square deviation (RMSD) for ensemble groups using different configurations of albedo (A), drainage (D) and stability adjustment (S) options. Scatterplots for runoff RMSD are shown in the upper right of the matrix and scatterplots for SCA RMSD are shown in the lower left. RMSD is calculated following the description in Section 4.3.1. Each point represents the RMSD for a single year, and results from all three climate input strategies are plotted (Section 3.2.3 and Section S1).

S4 Runoff and LST Anomaly Time Series

Figure S3 compares simulated catchment-scale snowpack runoff and LST monthly anomalies with observed runoff and MODIS remote sensing, respectively. This demonstrates that the simulated anomalies are reasonably consistent with the reference datasets in anomaly space. The spread amongst the four major ensemble groups considered in Section 4.3.2 is also fairly small relative to the amplitude of inter-annual variability. Some discrepancies are of course evident, which may partly reflect the limitations of the HAR climate product in capturing the sequencing of inter-annual climate anomalies (Section

4.3.2). Indeed, some of the discrepancies likely fit with the SCA errors discussed in Table 3. Overall then, Figure S3 supports the finding in Section 4.3.2 that the FSM ensemble may be somewhat reliable in anomaly space when climate input anomalies are sufficiently represented.
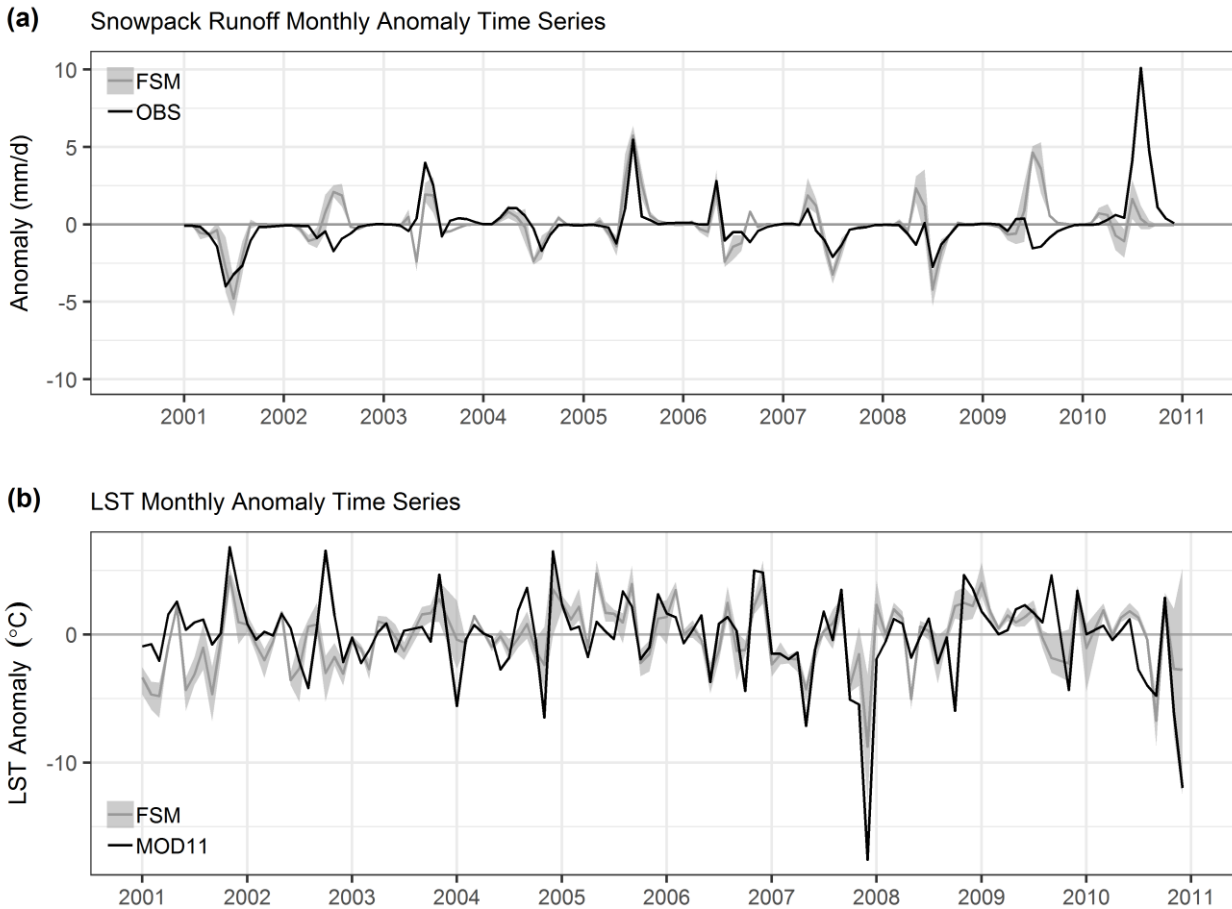
Figure S3. Comparison of simulated (a) runoff and (b) LST catchment-scale monthly anomalies with observations and remote sensing, respectively. The mean and range of the four primary FSM ensemble groups considered in Section 4.3.2 are shown by the grey lines and shading. The observations in (a) are total catchment runoff anomalies, while the simulated values are for snowpack runoff (Section 3.3).

S5 Climate Input Sensitivity

10   Figure S4 and Figure S5 show the implications of using the two alternative climate input strategies described in Section 3.2.3 and Section S1. These strategies are essentially HAR-based inputs with no bias correction (Figure S4), and using local observations to derive climate fields as far as possible (Figure S5). Figure S4 indicates that omitting bias correction of HAR temperatures leads to a very similar ensemble structure to that presented for the baseline climate inputs in Section 4.1. There is a small shift of the cumulative snowpack runoff curves to the right in Figure S4 compared with Figure 2, which reflects

15   the effect of the cold temperature bias in delaying runoff. However, the structure of ensemble groups matches closely in the two figures. Therefore temperature bias correction does not fundamentally alter FSM response in this case.

Figure S5 also shows notable similarity with Figure 2 in Section 4.1 in terms of overall ensemble structure. However, there is more spread within the principal groups within the ensemble in Figure S5, especially for slow-responding combinations (prognostic albedo and a representation of liquid water refreezing, retention and drainage). This leads to a

wider overall ensemble spread when applying primarily observation-based inputs. Nevertheless, the rank order of the primary and secondary groups within the ensemble is the same as for the baseline climate inputs. This strengthens the notion that the findings in Section 4 are likely to be similar when using various commonly applied climate input strategies.
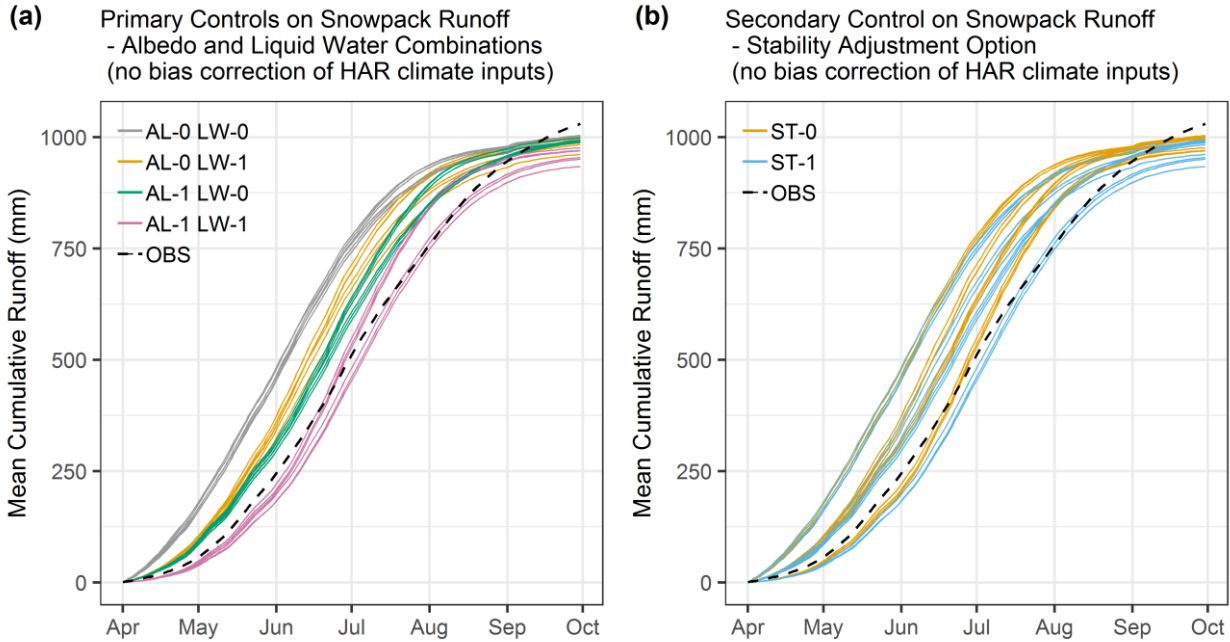
Figure S4. Comparison of mean cumulative snowpack runoff for the high-flow season for each of the 32 ensemble members with observed total runoff (OBS, black dashed line). In (a) each ensemble member is coloured according to the combination of albedo (AL) and liquid water (LW) parameterisations it uses. In (b) each ensemble member is coloured by its stability adjustment (ST) option. The results are based on input strategy (2) described in Section S1 (i.e. HAR-based but without temperature bias correction).
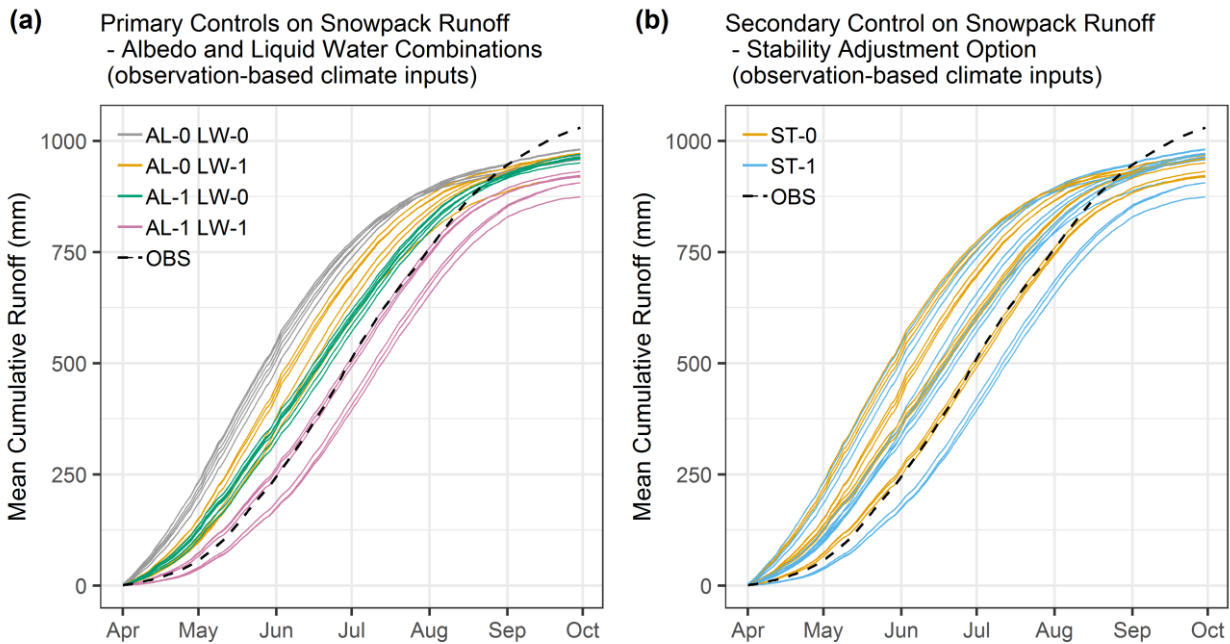
Figure S5. As Figure S4 but for input strategy (3) described in Section S1 (i.e. observation-based as far as possible).

References

Liston, G. E. and Elder, K.: A Meteorological Distribution System for High-Resolution Terrestrial Modeling (MicroMet), J. Hydrometeorol., 7, 217–234, doi:10.1175/JHM486.1, 2006.

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, 1910–1927, doi:10.1175/JCLI-D-13-00282.1, 2014.

Pellicciotti, F., Raschle, T., Huerlimann, T., Carenzo, M. and Burlando, P.: Transmission of solar radiation through clouds on melting glaciers: a comparison of parameterizations and their impact on melt modelling, J. Glaciol., 57(202), 367–381, doi:https://doi.org/10.3189/002214311796406013, 2011.

Pritchard, D. M. W., Forsythe, N., Fowler, H. J., O'Donnell, G. M. and Li, X.-F.: Evaluation of Upper Indus Near-Surface Climate Representation by WRF in the High Asia Refined Analysis, J. Hydrometeorol., 20(3), 467–487, doi:10.1175/JHM-D-18-0030.1, 2019.