Journal of Geophysical Research: Atmospheres

Supporting Information for

The energy and mass balance of Peruvian glaciers

Catriona L. Fyffe[1], Emily Potter[2], Stefan Fugger[3,4], Andrew Orr[5], Simone Fatichi[6], Edwin Loarte[7], Katy Medina[7], Robert Å. Hellström[8], Maud Bernat[3,9], Caroline Aubry-Wake[10], Wolfgang Gurgiser[11], L. Baker Perry[12], Wilson Suarez[13] Duncan J. Quincey[2], Francesca Pellicciotti[1,3]

[1]Engineering and Environment, Northumbria University, Newcastle-Upon-Tyne, UK

[2]School of Geography, University of Leeds, Leeds, UK

[3]Swiss Federal Institute for Forest, Snow and Landscape Research, Zürich, Switzerland

[4]Institute of Environmental Engineering, ETH Zürich, Zürich, Switzerland

[5]British Antarctic Survey, Cambridge, UK

[6]Department of Civil and Environmental Engineering, National University of Singapore, Singapore

[7]Instituto Nacional de Investigación en Glaciares y Ecosistemas de Montaña, Huaraz, Perú

[8]Bridgewater State University, Bridgewater, Massachusetts, USA

[9]Earth Sciences and Environment, Mines Paris Tech, Paris, France

[10]University of Saskatchewan, Saskatchewan, Canada

[11]Department of Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck, Austria

[12]Appalachian State University, Boone, USA

[13]Servicio Nacional de Meteorología e Hidrología del Perú, Perú

Contents of this file

Introduction

This document includes supplementary information on the WRF climate model methodology, the derivation of snowlines, details of the meteorological stations, the meteorological data cleaning and filling steps and full details of the Tethys-Chloris model used, including the Monte-Carlo uncertainty assessment. It also includes extended results, including the validation of the model and supplementary figures to support the results. Furthermore, it includes a table supporting the discussion of Peruvian glaciers in a South American context.

Text S1. Regional Climate and Glaciology

Text S1.1 WRF methodology

The Weather Research and Forecasting (WRF) model version 3.8.1 (Skamarock et al., 2008) was run from 1980 to 2018 over an outer domain with 12 km horizontal resolution covering Peru, and two inner domains with 4 km horizontal resolution covering the two study catchments (see main manuscript). For a full description of the WRF model physics and dynamics options, see Table S 1.

The precipitation and temperature data were bias-corrected against in-situ precipitation and air temperature observations in both regions. Thirty-five precipitation stations, and 26 temperature stations were used, from SENAMHI (Hunziker et al., 2017), ANA, ANTAMINA and CIAD. The data cover various years from 1980 to 2018, and all data were cleaned before use. For the temperature data, data were excluded where they were above 40 °C or below -20 °C, above or below three standard deviations from the mean, jumped by 10 °C and back again on consecutive days, checked where there were more than 5 consecutive days of the same temperature, and excluded where the minimum daily temperature exceeded the maximum daily temperature. Data were then checked again using the test of three standard deviations from the mean, as at some stations the erroneous data had skewed the mean substantially. Precipitation data were checked using a double mass analysis, where the data from one station is compared to all others, to check for jumps or changes to the gradient. Lastly, all data were examined by eye, and those stations which showed a trend were checked against nearby stations for differences. The model output is interpolated to the station locations using nearest-neighbour interpolation. Before bias-correction, the model air temperature at 2 m is adjusted for the discrepancy between the model elevation and the station elevation, using a temperature gradient derived from the model output. In general, the WRF output overestimates precipitation, both in the number of days on which precipitation falls ('wet days'), and in magnitude. In the wet season, the magnitude of the raw WRF data is

approximately double that of the observations, and the percentage bias is very variable in the dry season (when there is very little precipitation). The WRF output somewhat underestimates maximum daily temperature and does not adequately represent the annual cycle in minimum daily temperature.

For the bias correction, the model precipitation output is first corrected for the number of wet days, then a constant multiplicative correction is applied to adjust the magnitude. Furthermore, the model minimum and maximum daily temperatures are corrected additively, with the correction amount constant across the domain for the maximum daily temperature, and variable with latitude, longitude and height for minimum daily temperature.

Table S 1 Details of the Weather Research and Forecasting Model setup.

Domains and forcing data	
Number of domains	3
Horizontal grid resolution	12, 4, 4 km
Number of vertical levels	35
Model top	50 hPa
Topography data	Domain 1: U.S. Geological Survey 30 s; domains 2 & 3: Shuttle Radar Topography Mission (Jarvis et al., 2008)
Land surface and snow and ice data	Domain 1: U.S. Geological Survey 30 s; domains 2 & 3: U.S. Geological Survey 30 s, adjusted using the Randolph Glacier Inventory (Pfeffer et al., 2014)
Lateral boundary forcing data	ERA5, every 6 hours (Hersbach et al., 2020)
Spin-up period	Each year was run separately, with 1 month spin-up
Nesting	One-way nesting
Nudging	Spectral nudging above model level 15
Physics schemes	
Microphysics	Morrison double moment (Morrison et al., 2005)
Radiation	CAM scheme (Collins et al., 2004)
Surface layer	Revised MM5 (Jiménez et al., 2012)
Land surface	Noah-MP (multi-physics) (Niu et al., 2011)
Planetary boundary layer	Mellor-Yamada Nakanishi and Niino Level 2.5 (Nakanishi & Niino, 2004)
Cumulus	Kain-Fritsch scheme (Ma & Tan, 2009) in domain 1, none in domains 2-3
Sea surface temperature update	On
Heat and moisture fluxes from the surface	On
Snow cover effects	On
Cloud effects	On
Dynamics	
Diffusion	Calculated in real space
Eddy diffusion coefficient	Diagnosed from horizontal diffusion
Short-wave numerical noise filter	On
Top of model damping	Rayleigh damping in top 5,000 m of model
Time off-centering for vertical sound waves	0.5

Text S1.2 Deriving snow line altitudes

Top-of-atmosphere (TOA) scenes are selected in Google Earth Engine (GEE) for each sensor according to the period and catchment under study. We consider only scenes that exhibit

more than 50% cloud free area, merge same-date acquisitions from the same sensor, and clip the resulting data to each catchment. For each date of observation, the following methodology is applied:

First, all clouds and shadows are masked, as these can prevent effective identification of snow. We perform this masking for Landsat sensors using the QA band, which contains bitwise masks for cloud and shadow. For Sentinel-2, we also use the QA band for cloud masking, but the Hollstein method (Hollstein et al., 2016) proved to be more efficient for shadow masking. We additionally mask glaciers and persistent water surfaces, which can confound automated methods to identify snow cover, based on the Randolph Glacier Inventory 6.0 (Pfeffer et al., 2014) and the GlobalSurfaceWater dataset from Pekel et al., (2016).

From the remaining observable area, we then identify snow-covered area, using the methods of Girona-Mata et al. (2019), but implemented in GEE. We use the snow-covered area maps to derive a snow-cover frequency map for a requisite period (e.g. monthly, seasonally, or annually). Specifically, we calculate the snow-cover frequency for each pixel as the ratio of snow observations to total observations (e.g. neglecting the times the pixel was masked). We perform this analysis at 30 m spatial resolution and combine all sensors' records. Last, we use the spatial snow frequency results to describe the frequency distribution per elevation for the requisite period. The snowline elevations for the catchments are taken as the elevation of the 50^{th} centiles (Figure S1).

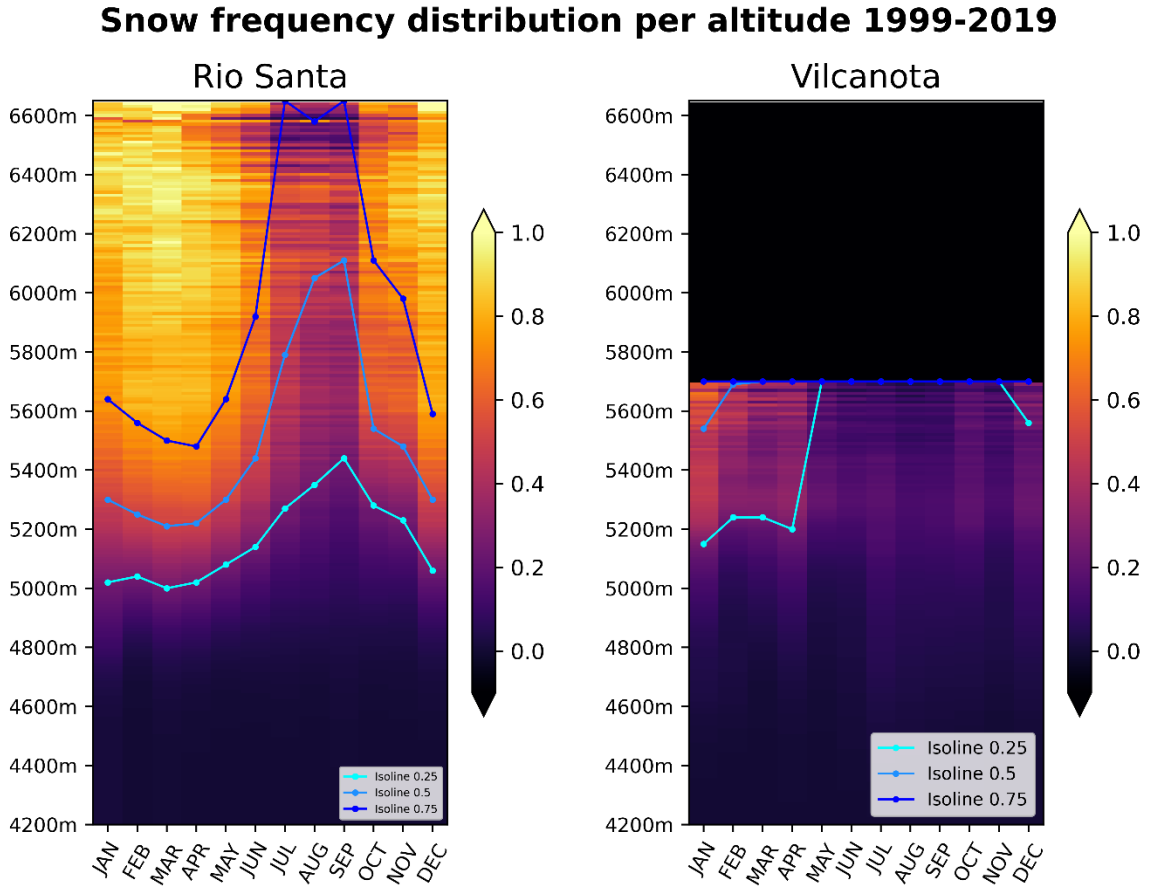

Figure S 1 Snow frequency distribution derived from the snow area detected on 1345 (Rio Santa) and 996 (Vilcanota) scenes from Landsat 5, 7, 8 and Sentinel 2 over the period 1999-2019. Dash lines corresponds to the 25th, 50th and 75th centiles. HR = Harmonic Regression of order 2 calculated on the isolines.

Table S 2 Glaciological characteristics of the two catchments and each of the study glaciers. The INAIGEM 2018 Glacier inventory was used to derive the glacier outlines and the elevation data are from ASTER GDEM v3 at 31 m horizontal resolution. Snowlines are derived for each region, excluding glaciated areas, as described above.

Glacier or region	Min. elevation (m a.s.l.)	Median elevation (m a.s.l.)	Max. elevation (m a.s.l.)	Station elevation (m a.s.l.)	Annual snowline (m a.s.l.)	Dry season snowline (m a.s.l.)	Wet season snowline (m a.s.l.)
Rio Santa catchment	4262	5211	6752	-	5500	5731	5273
Vilcanota catchment	4399	5290	6333	-	5687	5701	5674
Shallap Glacier (SG)	4651	5119	5946	4790	-	-	-
Artesonraju Glacier (AG)	4721	5146	5889	4797	-	-	-
Cuchillacocha Glacier (CG)	4729	5288	6031	4821	-	-	-
Quisoquipina Glacier (QQG)	5054	5331	5608	5180	-	-	-
Quelccaya Ice Cap (QIC)	4901	5442	5670	5650	-	-	-

Text S2. Study sites and data

Text S2.1 Study sites

Table S 3 Details of meteorological stations used in this study. Instrument heights are given for on-glacier stations used for modelling. For station type T = tripod on ice surface, D = drilled into ice/snow surface. *On Cuchillacocha Glacier the wind speed was measured above the other instruments, but it was corrected to a height of 1.32 m presuming a logarithmic wind speed profile. **The dates of the modelling period cover the time span modelled and so are not necessarily the same as the length of the record from the on-glacier stations; data filling from other stations was often required to complete the timeseries. Ta is air temperature, RH is relative humidity, $S\downarrow$ is incoming shortwave radiation, $S\uparrow$ is outgoing shortwave radiation, $L\downarrow$ is incoming longwave radiation, $L\uparrow$ is outgoing longwave radiation, Pre is air pressure, SR50 is a snow depth sensor and Rn is net radiation.

Station	Acronym	On or off glacier	Latitude (°)	Longitude (°)	Elevation (m a.s.l.)	Variables measured	Station type / instrument heights	Modelled period**
Artesonraju Carac Moraine	ACM	off	-8.9741	-77.6570	5050	Ta, RH, $S\downarrow$, WS	-	-
Artesonraju Glacier	AG	on	-8.9648	-77.6357	4797	Ta, RH, WS, $S\downarrow$, $S\uparrow$, $L\downarrow$, $L\uparrow$, SR50	T / 2 m	20/05/2006 – 12/05/2013
Artesonraju Moraine	AM	off	-8.9692	-77.6378	4817	Ta, RH, WS, Pr	-	-
Cuchillacocha Glacier	CG	on	-9.4054	-77.3521	4821	Ta, Ts, RH, WS, $S\downarrow$, $S\uparrow$, Rn	T / 1.32 m*	24/06/2014-05/08/2018
Cuchillacocha Quilcay	CQ	off	-9.4139	-77.3549	4642	Ta, RH, WS, Pr, $S\downarrow$,	2.00 m	-
Morder	Morder	off	-8.9633	-77.6451	4900	Pr	-	-
Santiago Antúnez de Mayolo	Santiago	off	-9.5165	-77.5249	3079	Pr	-	-
Shallap Glacier	SG	on	-9.4892	-77.3380	4790	Ta, RH, WS, Pre, $S\downarrow$, $S\uparrow$, $L\downarrow$, $L\uparrow$, SR50	D (before 24/06/2011) / variable, derived from SR50 and T / 2 m.	26/07/2010 – 30/11/2011 and 28/05/2012 – 18/09/2012
Shallap Moraine New	SMN	off	-9.4911	-77.3457	4767	Ta, RH, WS, $S\downarrow$, $S\uparrow$, $L\downarrow$, $L\uparrow$	-	-
Shallap Moraine Old	SMO	off	-9.4920	-77.3369	4945	Ta, RH, WS, Pr, $S\downarrow$	-	-
Quelccaya Ice Cap	QIC	on	-13.9197	-70.8165	5650	Ta, RH, WS, Pr, $S\downarrow$, SR50	D / variable, derived from SR50	17/07/2016 – 31/12/2018
Quisoquipina Glacier	QQG	on	-13.7944	-70.8852	5180	Ta, RH, WS, $S\downarrow$, $S\uparrow$, $L\downarrow$, $L\uparrow$	T / 2.5 m	27/10/2011-25/08/2016

Table S 4 Instruments installed on meteorological stations. All Ta and RH sensors were situated in a radiation shield, with (V) symbolising that Ta or RH measurements were artificially ventilated. Additional sensors and variables were measured at some stations, but for brevity only those used in modelling are included here. *Rn (net radiation) was measured at CG. Instrument precision is given in brackets.

Station	**T_a**	**RH**	**WS**	**$S\downarrow$**	**$S\uparrow$**	**$L\downarrow$**	**Pr**	**SR50**
ACM	Vaisala HMP45 (V) (±0.2°C)	Vaisala HMP45 (V) (±2%)	Young 05103 (±0.3 ms^{-1} or 1%)	Star pyranometer (0.3 to 3 µm)	-	-	-	-
AG	Vaisala HMP45 (±0.2°C)	Vaisala HMP45 (±2%)	Young 05103 (±0.3 ms^{-1} or 1%)	Kipp&Zonen CNR4 (5 to 20 µV W^{-1} m^2)	Kipp&Zonen CNR4 (5 to 20 µV W^{-1} m^2)	Kipp&Zonen CNR4 (5 to 20 µV W^{-1} m^2)	-	Campbell SR50a (±1 cm or 0.4% of distance)
AM	Vaisala HMP45 (V) (±0.2°C)	Vaisala HMP45 (V) (±2%)	Young 05103 (±0.3 ms^{-1} or 1%)	Star pyranometer (0.3 to 3 µm)	-	-	Ott Pluvio (± 0.05 mm)	-
CG	Campbell CS215 (±0.4°C)	Campbell CS215 (±2%)	Young 03002 (±0.5 ms^{-1})	Apogee SP-230 (0.20 mV W^{-1} m^2)	Apogee SP-230 (0.20 mV W^{-1} m^2)	As derived from R_n*: Kipp&Zonen NR-Lite2 (10 mV W^{-1} m^2)	-	-
CQ	Onset HOBO Temperature RH Smart Sensor: S-THB-M002 (±0.2°C)	Onset HOBO Temperature RH Smart Sensor: S-THB-M002 (±2.5 %)	S-WSB-M003 (± 1.1 ms^{-1})	Apogee SP110 (±5%)	-	-	0.2 mm tipping bucket Smart Sensor S-RGB-M002 (±0.2 mm or ±1.0%). No wind shield	-
SG	Vaisala HMP45 (±0.2°C)	Vaisala HMP45 (±2%)	Young 05103 (±0.3 ms^{-1} or 1%)	Kipp&Zonen CNR4 (5 to 20 µV W^{-1} m^2)	Kipp&Zonen CNR4 (5 to 20 µV W^{-1} m^2)	Kipp&Zonen CNR4 (5 to 20 µV W^{-1} m^2)	-	Campbell SR50 (±1 cm or 0.4% of distance)
SMN	Vaisala HMP45 (±0.2°C)	Vaisala HMP45 (±2%)	Young 05103 (±0.3 ms^{-1} or 1%)	Kipp&Zonen CNR4 (5 to 20 µV W^{-1} m^2)	Kipp&Zonen CNR4 (5 to 20 µV W^{-1} m^2)	Kipp&Zonen CNR4 (5 to 20 µV W^{-1} m^2)	Ott Pluvio (± 0.05 mm). No wind shield.	-
SMO	Vaisala HMP45 (V) (±0.2°C)	Vaisala HMP45 (V) (±2%)	Young 05103 (±0.3 ms^{-1} or 1%)	Star pyranometer (0.3 to 3 µm)	-	-	Ott Pluvio (± 0.05 mm). No wind shield.	-
QIC	Campbell EE181 (±0.2°C)	Campbell EE181 (±1.3 + 0.003 • RH reading %)	Young 05103-45 (±0.3 ms^{-1} or 1%)	Apogee CS-300 (±5%)	-	-	Ott Pluvio 2. (± 0.05 mm) Double wind shield.	Campbell SR50a (±1 cm or 0.4% of distance)
QQG	Vaisala HMP45C (±0.2°C)	Vaisala HMP45C (±2%)	Young 05103 (±0.3 ms^{-1} or 1%)	Kipp&Zonen CNR1 (7 to 15 µV W^{-1} m^2)	Kipp&Zonen CNR1 (7 to 15 µV W^{-1} m^2)	Kipp&Zonen CNR1 (7 to 15 µV W^{-1} m^2)	-	-

Text S2.2 Data cleaning

Meteorological data were carefully cleaned and quality checked prior to further analysis. First clearly erroneous data were removed (deleted from the record), then specific corrections were applied to certain variables to ensure they remained within a certain range (Table S 5). Any removed data were filled where possible, using either data from nearby off glacier stations, data from other years from the same station, or WRF outputs, as outlined in Text S2.3 and Table S 6 for each variable and station. Any records of zero wind speed values were removed because although zero wind speed is possible for short periods, sometimes these periods were relatively long, often overnight and coincident with Ta<0°C. These zero values could be explained by the freezing of the anemometer in cold conditions (Oerlemans, 2010) or the accuracy of the anemometer to small wind speeds. Zero wind speeds are also problematic in energy balance modelling as the turbulent fluxes are reduced to zero. Gaps caused in the wind speed records by removing zero wind speeds were filled by replacing them with hourly average wind speed values for each month. Some variables were also checked for large jumps: air temperature for a change of more than 10°C and relative humidity for a change of more than 40%. They were used to identify areas of the record for inspection, with data removed if it were deemed erroneous. Similarly, if values were the same for six consecutive time steps (excluding shortwave radiation = 0 W m^{-2} and relative humidity = 100%) then the data were inspected and removed if necessary.

Table S 5 Corrections applied to meteorological variables. With RH = relative humidity, SW = shortwave radiation, $S\downarrow$ = incoming shortwave radiation, $S\uparrow$ = outgoing shortwave radiation and WS = wind speed.

Variable	**Cleaning rules**
RH	if RH>100%, then RH=100%
SW	if $S\downarrow$< 0 W m^{-2}, then $S\downarrow$= 0 W m^{-2}
	if $S\uparrow$< 0 W m^{-2}, then $S\uparrow$= 0 W m^{-2}
	if $S\uparrow > S\downarrow$, then $S\uparrow = S\downarrow$
WS	if WS≤ 0, then WS removed.

Text S2.3 Data corrections and filling steps

To create continuous timeseries for all five stations any gaps in the on-glacier data were filled. Where data were taken from an off-glacier station it was corrected to be representative of on-glacier conditions. Multiple methods were tested, including various linear regressions and the average hourly difference between the two stations. The aim was to maintain the diurnal and seasonal cycle at the on-glacier station. Full details of all data corrections and filling steps are given in Table S 6, although the filling regime will be briefly summarised here. Where there were no $S\uparrow$ or $L\downarrow$ data or gaps in measured timeseries, these were modelled using the parameterisations in Text S3. At Shallap all variables were measured, except Pr which was derived from Shallap Moraine Old and Santiago; at Artesonraju data were filled using relationships with Artesonraju Moraine and Artesonraju Carac Moraine (only $S\downarrow$) and where necessary filled from other years, except Pr which was derived from Artesonraju Moraine, Morder and WRF data; Cuchillacocha data were available

for only two weeks, so the record is predominantly derived from Cuchillacocha Quilcay; at Quelccaya most data were measured at Quelccaya, except $S\downarrow$ which was partly erroneous and filled with WRF; and at Quisoquipina most variables were measured at Quisoquipina, with some filling from another year, except Pr, which was entirely from WRF. An assessment of the suitability of using WRF data for Quisoquipina precipitation and Quelccaya incoming shortwave radiation is given Text S2.4.

Unfortunately, there were data gaps at all the Shallap stations for all variables and at all stations in December 2011 and for radiation and wind speed variables between 07/02/2012 and 20/04/2012. Given the relatively short timeseries at Shallap Glacier, and the lack of overlap of Shallap Glacier with Shallap Moraine New and Shallap Moraine Old for $S\downarrow$, it was deemed not appropriate to fill these gaps with data from other years. The modelled timeseries therefore has a gap from 1/12/2011 00:00 until 27/05/2012 23:00. The end and beginning of the gap were chosen to coincide with snow free conditions. Creating the on-glacier record at Cuchillacocha Glacier was complicated by the on-glacier data only overlapping with the off-glacier data for 339 hours (just over 14 days) from the 26th June 2014 until the 10th of July 2014. This means that the relationships between the on and off glacier data are based on the winter (dry season) conditions only, although for Ta, $S\downarrow$ and WS the earlier sunrise and longer day length were accounted for.

Text S2.4 Assessment of the suitability of using WRF data at Quisoquipina and Quelccaya

When using WRF data to fill the on-glacier records data were extracted for the relevant grid cell of WRF. Where WRF data composed the majority of an input variable we conducted extra checks of its suitability.

To check that the WRF precipitation modelled for Quisoquipina (QQG) was appropriate it was compared to nearby measured precipitation records, including Quelccaya (QIC) and two records from Sibinacocha dam, one from SENAMHI (SC) with an elevation of 4880 m a.s.l, and another (SCD) with an elevation of 4895 m a.s.l. The precipitation was summed for the time period when all three stations had data (from 18/07/2017 12:00 until 31/12/2018 at 23:00). Figure S 2 shows the comparison of the precipitation sums for all locations and the QQG WRF precipitation. The precipitation sum at QQG from the linear relationship (based only on the measured stations) equates to 1352 mm, which is 112 mm or 8.3% less than that given by WRF at QQG (1464 mm). The difference is less than the difference in precipitation between the two stations at Sibinacocha dam (125 mm). It was not possible to use data directly from these stations with a lapse rate applied for the QQG modelling, as measured data from any of the stations only overlapped a small proportion of the QQG modelling period. The WRF data for QQG was therefore used directly for modelling.

To check the suitability of WRF incoming shortwave radiation modelled for QIC, it can be compared to the data measured at QIC. The mean error (QIC – WRF) equals to –64 W m^{-2}, and the RMSE equals 160 W m^{-2}. The negative mean error is likely because the errors in the

QIC data were often around midday, resulting in gaps in the diurnal cycle. Although extra cleaning methods were applied to the QIC data (any days with < 12 hours of positive incoming shortwave radiation were removed, and any data between 07:00 and 17:00 which equalled zero were removed) any remaining errors likely underestimate incoming shortwave radiation. Some of the overestimation of incoming shortwave radiation by WRF may be real but given the gaps in the QIC data this was difficult to ascertain and therefore the WRF data were used when the QIC data were missing.

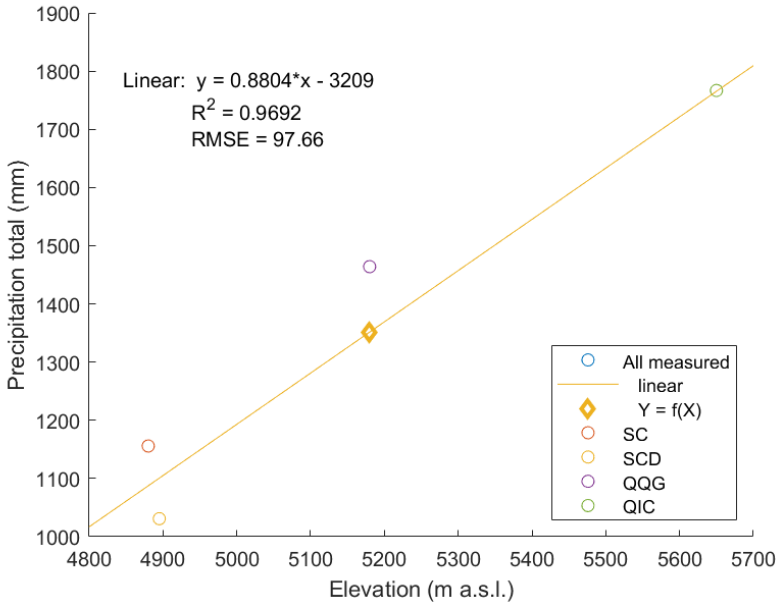

Figure S 2 Comparison of precipitation sums for Sibinacocha (SC), Sibinacocha Dam (SCD), QQG and QIC over the period 18/07/2017 12:00 until 31/12/2018 23:00. The linear relationship was created with the points from stations with measured data (SC, SCD and QIC). The diamond gives the estimated precipitation sum from this relationship at the elevation of QQG.

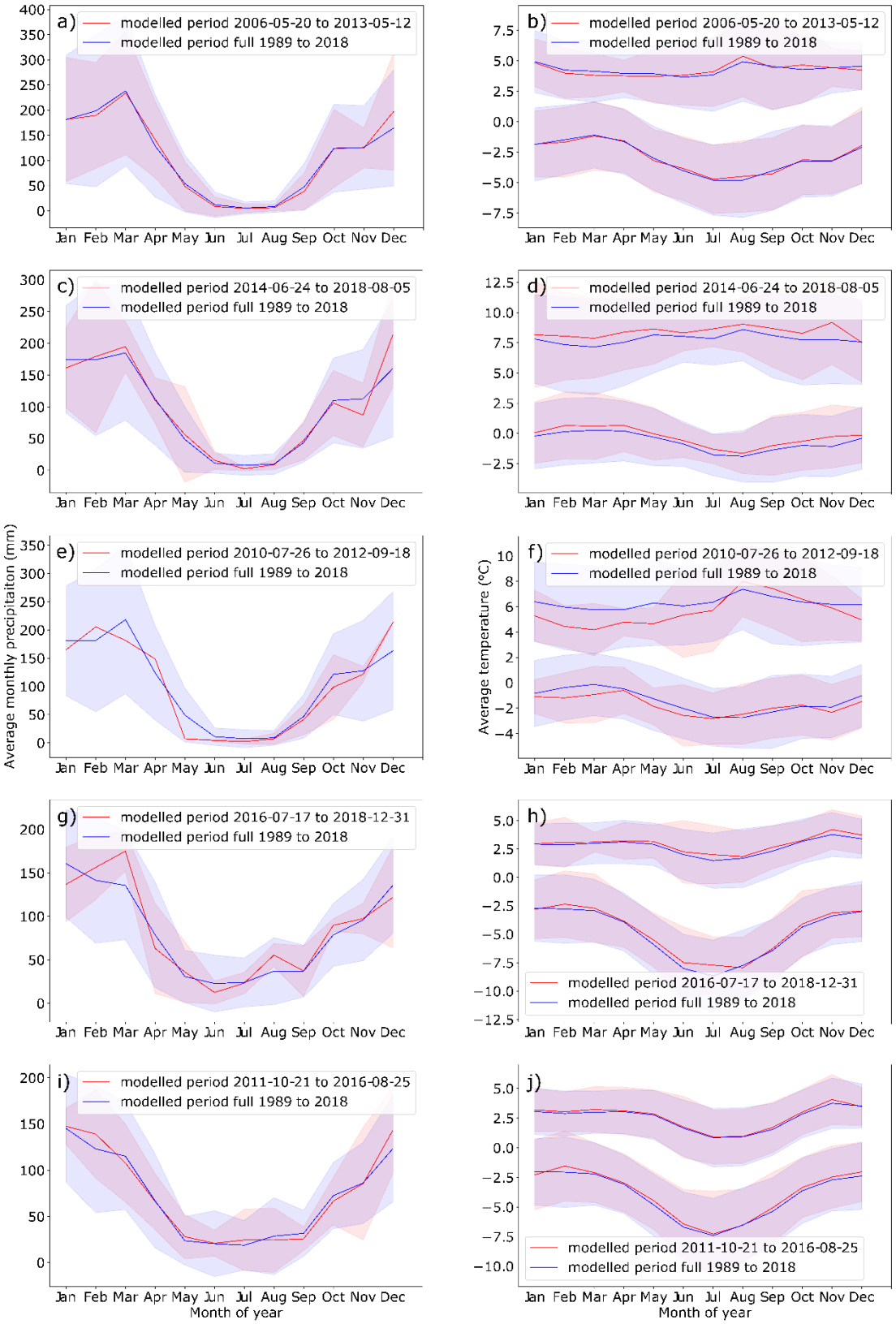

Figure S 3 The average annual cycle of monthly precipitation, from the WRF model, over the period where observations are available (red) and over the full period 1980-2018 (blue) from the locations of the weather stations at a) Artesonraju Glacier, c) Cuchillacocha Glacier, e) Shallap Glacier, g) Quelccaya Ice Cap, i) Quisoquipina Glacier. The annual cycle of monthly averaged minimum and maximum daily temperatures are shown in figures b) to j). Shading represents two standard deviations from the mean.

Table S 6 Details of data correction and filling steps for each meteorological station and input variable. Please see the main text for the abbreviations of the station names and variable definitions, with RH relative humidity and WS wind speed. Data are defined as on-glacier (for the percent on-glacier column) if they were measured at the on-glacier station, including where small gaps (less than ~ 1 day) were filled with the on-glacier data and including where zero wind speed data were filled with the monthly diurnal average, but not including where data were filled over longer periods (for instance using the same day from a different year). Precipitation from the 'main' nearby off-glacier site is also included as 'on-glacier' for this purpose.

Station	**Donor stations**	**Variable**	**Corrections and filling steps**	**Percent on-glacier**
SG	SMO, SMN, Santiago	T_a	As measured at station. One small gap filled with interpolation, one day gap filled with average of the days before and after.	100
		RH	As measured at station. One small gap filled with interpolation, one day gap filled with average of the days before and after.	100
		WS	As measured at station. Zero values removed and replaced with hourly average value for each month. Small gaps also filled in this way.	100
		$S\downarrow$	As measured at station. One small gap filled with interpolation, one day gap filled with average of the days before and after.	100
		$S\uparrow$	As measured at station. One small gap filled with interpolation, one day gap filled with average of the days before and after.	100
		$L\downarrow$	As measured at station. One small gap filled with interpolation, one day gap filled with average of the days before and after.	100
		Pr	No Pr data at SG. Used SMO when available. When gaps in SMO used data from Santiago precipitation station. The Santiago station is at a much lower elevation than SMO so precipitation amounts were corrected by finding the daily magnitude of precipitation that was required to be added to the Santiago precipitation in order that the precipitation sum for SMO and Santiago matched. Due to precipitation occurring less frequently at Santiago than SMO the correction likely overestimates the precipitation lapse rate for a given hour but it accounts for the decrease in precipitation frequency at Santiago compared to SMO. Santiago precipitation is available only as daily sums so it was converted to hourly by applying the average hourly precipitation ratio (in other words the proportion of precipitation that falls in each hour compared to the daily total) derived for SMO. In April 2011 there were gaps remaining, these were filled with SMO data from the same day of a randomly chosen year (between 2005 and 2012). The SMO precipitation gauge is an OTT Pluvio with no wind shield and the gauge type at Santiago is unknown. Therefore an undercatch ratio of 0.63 (as given at the CARE site within Nitu et al. (2018) for a Pluvio gauge with no wind shield under snow conditions) was applied to the whole record when precipitation was	26

			modelled as snow based on the Ding et al. (2014) precipitation partition algorithm. The corrected snowfall was added to the liquid precipitation to give the final corrected timeseries.	
AG	ACM, AM, Morder	T_a	Gaps in AG filled with AM, based on the combination of a 2nd order polynomial relationship (R^2 = 0.811) when AM T_a < 6.82°C, and a linear relationship at higher temperatures (R^2 = 0.476). The RMSE of T_a derived from AM = 0.80°C. Small number of remaining gaps filled with the same day of a randomly chosen year (2005-2012).	32
		RH	Gaps in AG filled with AM, based on a third order polynomial (R^2 = 0.915). RMSE of RH derived from AM = 5.38%. Remaining gaps filled with the same day from a randomly chosen year between 2005 and 2012.	32
		WS	Zero values removed and replaced with hourly average value for the calendar month. Gaps in AG filled with AM data corrected to AG based on the hourly average difference between AG and AM. The hourly average difference was calculated (and applied) to each season separately (winter months are June to Sep, summer months are Oct to May). Where the derived wind speed was less than zero the value from AM was used directly. Gaps in AM were filled with the same day from a different and randomly chosen year (from 2005 to 2012) and then the average hourly difference between AG and AM calculated from the data above was applied to give the full timeseries. The RMSE between the derived wind speed data and AG = 2.16 ms^{-1}.	36
		$S\downarrow$	Gaps filled initially with hourly average relationship with ACM. Any derived values less than 0 are given a value of 0. The RMSE between AG and the series derived from ACM = 107.2 W m^{-2}. Remaining gaps (no AG nor ACM data) filled with AG data from the same day of a randomly chosen year.	92
		$S\uparrow$	AG data used where available, otherwise modelled within T&C.	92
		$L\downarrow$	AG data used where available, otherwise modelled within T&C.	92
		Pr	No Pr data at AG. AM data used instead when available. When not available, used data from the precipitation station Morder. The Morder data is given as a daily sum, so it was adjusted to hourly using the average hourly precipitation ratio (in other words the proportion of precipitation that falls in each hour compared to the daily total) derived for AM. Remaining gaps were filled with WRF modelled precipitation data. The AM precipitation gauge is an OTT Pluvio weighing gauge with no wind shield, and the gauge type at Morder is not known. Therefore an undercatch ratio of 0.63 (as given at the CARE site within Nitu et al. (2018) for a Pluvio gauge with no wind shield under snow conditions) was applied to the whole record when precipitation was modelled as snow based on the Ding et al. (2014) precipitation partition algorithm. The corrected snowfall was added to the liquid precipitation to give the final corrected timeseries.	83
CG	CQ	T_a	Just over 2 days of data from on-glacier station (CG), rest derived from relationship with off-glacier station (CQ). Note that there were two weeks of on-glacier data but the model was run with the data derived from CQ as far as possible to give a fair validation of the model against the stake validation data. Relationship	0.3

			with CQ is quadratic and was obtained without data at 9, 10 and 11am (R^2 (Adj) = 0.893), until Ta at CQ = 7.621°C, when a linear relationship is applied (R^2(Adj) = 0.137) (also without data from 9, 10 and 11am). At 9, 10 and 11am in winter (May to August) Ta is derived from CG Ta at 8am plus the difference in CG average hourly temperature between each hour. In the summer (September to April) Ta at 8 and 9am is derived from Ta at 7am using the same corrections as before but given to the hour before. The RMSE between the timeseries derived from CQ using the above method and the measured CG data is 0.65°C.	
		RH	Just over 2 days of data from on-glacier station (CG), rest derived from relationship with off-glacier station (CQ). Applied the average hourly difference between CQ and CG to the CQ record, the RMSE between CG and the derived record when data are available is 5.5%. In the summer the average hourly correction applied is moved forward 1 hour. Any values in the derived time series greater than 100% are corrected to a value of 100%.	0.3
		WS	Just over 2 days of data from on-glacier station (CG), rest derived from relationship with off glacier station (CQ). Zero wind speeds at CQ were removed and replaced with the average hourly cycle for that month. The CG wind speeds were derived from the average hourly difference between CG and CQ (RMSE = 0.613). The average hourly differences were applied as found for the winter months but were brought forward by one hour in the summer months.	0.3
		$S\downarrow$	Just over 2 days of data from on-glacier station (CG), rest derived from relationship with off glacier station (CQ). In general, applied the average hourly ratio of CQ to CG (RMSE = 153 W m^2). In the winter (March to August) the ratios are applied as derived from the data. However, in the summer (September to February) the ratios are moved forward by one hour and applied between 00:00 and 10:00. To account for the longer summer day length, the ratios applied between 11:00 and 23:00 are themselves modelled from a quadratic relationship (R^2 (Adj) = 0.924) between the CG/CQ average hourly ratios and the hourly average CQ $S\downarrow$. This then allows the summer ratios to be found from the average hourly summer $S\downarrow$ at CQ, these can then be applied to give the derived CG values.	0.3
		$S\uparrow$	Modelled from albedo derived within T&C.	0
		$L\downarrow$	Modelled from cloudiness derived from $S\downarrow$ within T&C.	0

		Pr	Not measured at CG so used CQ data. CQ Pr is from a non-heated and non-shielded tipping bucket gauge, this resulted in a spike in the average hourly precipitation between 08:00 and 12:00 due to melt of snow that had accumulated in the gauge. The average hourly precipitation was calculated, and the corrected values for 8am to 12am found by interpolation between 7am and 1pm. The precipitation ratio between the original and adjusted average hourly pattern was then then used to remove a proportion of all precipitation between 8am and 12am. The 'removed' precipitation was added 7 hours earlier between 1am and 5am when T_a at CG was likely below 0°C. To correct for undercatch due to the lack of a wind shield the undercatch correction derived by Chubb et al. (2015) for an unshielded tipping bucket rain gauge during snow conditions was applied. This correction requires the precipitation and wind speed measured at CQ and is applied when snow was predicted to occur at the CQ site based on the Ding et al. (2014) precipitation partition algorithm. The original sum of snowfall at CQ was = 1131 mm, and after adjustment = 1990 mm, giving a snow undercatch ratio of 0.57. The corrected snowfall was added to the liquid precipitation to give the final corrected timeseries.	100
QIC	-	T_a	Measured at station.	100
		RH	Measured at station.	100
		WS	Measured at station. Zero values removed and replaced with hourly average value for the calendar month.	100
		$S\downarrow$	Some data available at station, but often erroneous. Remainder filled with WRF modelled inputs.	51
		$S\uparrow$	Modelled from albedo derived within T&C.	0
		$L\downarrow$	Modelled from cloudiness derived from $S\downarrow$ within T&C.	0
		Pr	Measured at station. Weighing gauge with double wind shield so no need for undercatch corrections. Pr corrected by 4% to take account of the gravity anomaly.	100
QQG	-	T_a	Measured at station, gaps filled with data on same date of randomly chosen year.	99
		RH	Measured at station, gaps filled with data on same date of randomly chosen year.	99
		WS	Measured at station. Zero values removed and replaced with hourly average value for the calendar month.	100
		$S\downarrow$	As measured at station, one timestep filled with interpolation.	100
		$S\uparrow$	As measured at station, one timestep filled with interpolation.	100
		$L\downarrow$	Measured at station.	100
		Pr	Modelled by WRF.	0

Text S3. Model description and methods

The overview of the Tethys-Chloris (T&C) model approach is given in the paper, however here we outline the approach taken to calculate each of the individual fluxes and determine the overall mass balance.

Text S3.1 Radiative fluxes

R_n is given by:

$$R_n = S\downarrow(1-\alpha_i) + L\downarrow + L\uparrow,$$

With $S\downarrow$ incoming shortwave radiation at the measurement height, α_i the snow or ice albedo, $L\downarrow$ incoming longwave radiation and $L\uparrow$ outgoing longwave radiation. Measured hourly albedo derived from $S\uparrow/S\downarrow$ is used for α_i for both snow and ice surfaces, where measurements of $S\uparrow$ (outgoing shortwave radiation) are available. Otherwise, α_i is modelled based on the surface type: for ice a single value is used (derived from measurements at the appropriate site); and for snow the Brock et al. (2000) snow albedo parameterisation is applied, including both the deep and shallow snow albedo equations. Since the Brock et al. (2000) parameterisation was developed at a daily scale and the modelling approach employed is hourly, a threshold precipitation value is added Pr_t, which must be exceeded by the sum of the precipitation over the previous 24 hours to refresh the snow albedo to that of new snow (typically Pr_t = 10 mm). We use the parameters as given in Brock et al. (2000) rather than modifying them for each site by calibration to measured albedo, since of the three sites to use modelled albedo (Artesonraju, Cuchillacocha and Quelccaya) only Artesonraju had sufficient data for this process. However, the rarity of the formation of a snowpack for long periods at Artesonraju meant the calibration tended towards extreme values which reduced the fit of the modelled melt to the validation data.

Values of $L\downarrow$ are derived from measurements when available and otherwise derived from cloudiness using:

$$L\downarrow = K\varepsilon_{cs}\sigma T_{aK}^4,$$

Where K is the attenuation cloud cover, ε_{cs} the clear sky emissivity, σ the Stefan-Boltzmann constant ($5.6704 \cdot 10^{-8}$ W m^{-2} K^{-4}) and T_{aK} the air temperature in Kelvins. ε_{cs} is derived following Dilley and O'Brian (1998):

$$\varepsilon_{cs} = \frac{a+b\left(\frac{T_{aK}}{273.16}\right)^6 + c\sqrt{\frac{w}{25}}}{\sigma T_{aK}^4},$$

Where $a = 59.38$, $b = 113.7$, $c = 96.96$ and w is precipitable water (kg m^{-2}) determined by the method of Prata (1996). K is derived following Unsworth and Monteith (1975):

$$K = N(1 - 0.84) + \frac{0.84N}{\varepsilon_{cs}},$$

with cloudiness (N) derived from the potential incoming shortwave radiation ($S_{pot}\downarrow$, W m^{-2}) using the common approach described by Juszak and Pellicciotti (2013):

$$N = 1 - \left(\frac{S\downarrow}{S\downarrow_{pot}}\right).$$

N can only be calculated during the day, defined as periods when $S\downarrow_{pot}$ is greater than a threshold (150 W m^{-2}), with nighttime values linearly interpolated between the daytime values. However, significant shading at Cuchillacocha during the morning resulted in erroneously high cloudiness values. To correct for this Cuchillacocha daytime was defined as being between 11:00 and 16:00, based on analysis of the diurnal $S\downarrow$ cycle. Timesteps with precipitation are always given a cloudiness of 1. Meanwhile $L\uparrow$ is computed from the Stefan-Boltzmann law:

$$L\uparrow = -\varepsilon\sigma T_{sK}^4,$$

With ε the surface emissivity (0.97 for both snow and ice) and T_{sK}, the surface temperature in Kelvins.

Text S3.2 Turbulent fluxes

H is calculated from:

$$H = \rho_a C_p \frac{(T_s - T_a)}{r_{ah}},$$

Where ρ_a is air density (kg m^{-3}), C_p is the specific heat capacity of air at constant pressure (J kg^{-1} K^{-1}), T_s is the temperature of the ice or snow surface (°C), T_a the air temperature (°C) at the measurement height and r_{ah} is the aerodynamic resistance to heat flux (s m^{-1}). The aerodynamic resistance is calculated using the simplified solution of the Monin-Obukohv similarity theory (Mascart et al., 1995; Noilhan and Mafhouf, 1996). Full details are given in the supplementary information of Fatichi et al. (2012a) and Fatichi (2010). The roughness lengths (m) of heat (z_{0h}) and water vapour (z_{0w}) used in the calculation of the aerodynamic resistance are equal in T&C ($z_{0h} = z_{0w}$), and $z_{0h} = z_{0w} = 0.1z_{0m}$, with the roughness length of momentum (z_{0m}) equal to 0.001 m for snow and ice surfaces. Roughness lengths obtained by calibration of the modelled latent heat flux against sublimation measured using lysimeters have been found on Artesonraju Glacier, Peru (without penitentes z_{0m} = 0.002 m, z_{0h}/z_{0w}, = 0.001 m, with penitentes z_{0m} = 0.02 m, z_{0h}/z_{0w} = 0.01 m (Winkler et al., 2009)), Zongo Glacier, Bolivia ($z_{0h}/z_{0w}/z_{0m}$ = 0.002 m to 0.03 m (Wagnon et al., 1999)) and Antizana Glacier 15, Ecuador ($z_{0h}/z_{0w}/z_{0m}$ = 0.0022 m to 0.0046 m (Favier et al., 2004a). These values would suggest that the z_{0m} value of 0.001 m is towards the low end of the range. However, Sicart et al., (2011) modelled the mass balance of Zongo Glacier using the relation z_{0h} = z_{0w} = z_{0m}/100, with z_{0m} = 0.01 m (based on wind and temperature profile data and eddy

covariance measurements), which results in the same z_{0h} and z_{0w} values as applied in this study.

λE is estimated from:

$$\lambda E = \lambda_s \frac{\rho_a(q_{sat}(T_s) - q_a)}{r_{aw}},$$

Where λ_s is the latent heat of sublimation defined as $\lambda_s = \lambda + \lambda_f$, with λ the latent heat of vapourisation (1000(2501.3-2.361T_a) J kg^{-1}) and λ_f the latent heat of melting (333700 J kg^{-1}). The term q_{sat} is the surface specific humidity at saturation, q_a is the specific humidity of air at the measurement height and r_{aw} the aerodynamic resistance to the vapour flux, which equals r_{ah}.

Text S3.3 Incoming heat with precipitation

To estimate Q_v the precipitation amount and temperature must be known. The rain/snow temperature is assumed to be the greater/smaller value of T_a and 0°C. Therefore Q_v is calculated as:

$$Q_v = c_w P_{r,liq} \rho_w [\max(T_a, 0) - T_s] + c_i P_{r,sno} \rho_w [\min(T_a, 0) - T_s],$$

Where c_w = 4186 J kg^{-1} K^{-1} is the specific heat of water, c_i = 2093 J kg^{-1} K^{-1} is the specific heat capacity of ice, ρ_w = 1000 kg m^{-2} is the density of water and $P_{r,liq}$ and $P_{r,sno}$ are liquid and solid precipitation (m s^{-1}).

Text S3.4 Ground heat flux

The definition of the ground heat flux G (W m^{-2}) differs slightly based on the surface type. In the case of snow cover, it is equal to the energy passed from the snowpack to the underlying ice surface:

$$G_{sno}(t) = \lambda_{sno} \frac{T_{sno}(t-1) - T_{ice}(t-1)}{d_{sno}},$$

where λ_{sno} (W m^{-1} K^{-1}) is the thermal conductivity of snow, T_{sno} (°C) is the surface temperature of the snow, T_{ice} (°C) is the surface temperature of the ice and d_{sno} (m) is the snow depth.

For ice under the absence of snow, G is the heat flux passed from the ice pack to the underlying surface or if the ice pack is greater than 2 m w. e. then to that depth within the ice:

$$G_{ice}(t) = \lambda_{ice} \frac{T_{ice}(t-1) - T_{grd}(t-1)}{d_{ice}},$$

where λ_{ice} (W m^{-1} K^{-1}) is the thermal conductivity of ice, T_{grd} (°C) is the temperature of the underlying layer, and d_{ice} (m) is the ice thickness.

Text S3.5 Snowpack water content

The water content of the snowpack is approximated using a bucket model, in which outflow of water from the snowpack occurs when the maximum holding capacity of the snowpack is exceeded. The maximum holding capacity of the snowpack is based on the snow water equivalent, holding capacity coefficient and snow density, following the method of Belair et al. (2003). Snowmelt plus liquid precipitation, minus the water released from the snowpack gives the current snowpack water content (Sp_{wc}). If the surface temperature of the snow T_{sno} is greater than 0°C then the snowpack water content is presumed to be liquid, whereas otherwise it is presumed frozen. The process of melting (resulting in a negative flux) and freezing (resulting in a positive flux) of the water content of the snowpack is associated with the heat flux Q_{fm}:

$$Q_{fm}(t) = \begin{cases} f_{sp}\frac{\lambda_f \rho_w Sp_{wc}(t-dt)}{1000\,dt}, & T_{sno}(t) < 0 \text{ and } T_{sno}(t-dt) \geq 0 \\ f_{sp} - \frac{\lambda_f \rho_w Sp_{wc}(t-dt)}{1000\,dt}, & T_{sno}(t) \geq 0 \text{ and } T_{sno}(t-dt) < 0 \end{cases},$$

Where dt (s) is the timestep and f_{sp} is the fraction of the snowpack water content involved in either melting or freezing. This fraction is defined using $f_{sp} = 5/WE_s$, with WE_s the water equivalent snowpack.

Text S3.6 Ice water content

The water content of ice is approximated with a linear reservoir model. The liquid water outflow is proportional to the ice pack water content (Ip_{wc}), which is initiated when the water content exceeds a threshold capacity, prescribed as 1% of the ice pack water equivalent. The icepack water content is the sum of ice melt and liquid precipitation, minus the water released from the ice pack. The water released is the sum of the ice pack excess water content plus the outflow from the linear reservoir, given as $I_{out} = \frac{Ip_{wc}}{K_{ice}}$, where $K_{ice}$is the reservoir constant which is proportional to the ice pack water equivalent. Unlike within snowpacks, Q_{fm} is not accounted for within the ice pack, since water is presumed to percolate quickly and avoid refreezing.

Text S3.7 Precipitation partition

Input precipitation is required to be partitioned into solid ($P_{r,sno}$) and liquid ($P_{r,liq}$) precipitation, because of the differing impacts of snow and rain on the energy and mass balance. Originally T&C employed the method of Wigmosta et al. (1994) where the precipitation is partitioned as a function of solely T_a, with all precipitation deemed snow under the condition $T_a \leq T_{min}$, all precipitation deemed rain under the condition $T_a \geq T_{max}$, solid and liquid precipitation partitioned between both states under intermediate conditions,

and T_{min} and T_{max} set as parameters. For this study, the precipitation partition method described by Ding et al. (2014) was additionally implemented in T&C. This scheme determines the precipitation partition based on the wet-bulb temperature, station elevation and vapour pressure. Ding et al. (2014) found that the wet-bulb temperature was found to be a better predictor than T_a of the precipitation type; that the temperature threshold between snow and rain is increased at higher elevations; and that the probability of sleet is reduced in conditions of low relative humidity.

To compare the effect of the precipitation partition on modelled snowfall and mass balance we ran the model at all sites using both approaches. At all sites there is an increase in the percentage of precipitation deemed snowfall using the Ding et al. (2014) approach, which consequently results in an increase in snow melt and a decrease in ice melt, since more snow accumulates and is available to melt (Table S 7). The exception to this is at Quelccaya, where snow melt is slightly decreased, here only snow melt is possible, so it is likely the increased rainfall under the Wigmosta et al. (1994) approach would both reduce the snow albedo (due to a less frequent refresh to the new snow albedo) and increase the heat flux due to precipitation, therefore increasing the energy for melt. At all sites ablation is reduced by the Ding et al. (2014) approach compared to the Wigmosta et al. (1994) approach, with the difference largest at Artesonraju. This analysis highlights the need to take account of the influence of the relative humidity and elevation in the partition of precipitation in the Peruvian Andes. The Ding et al. (2014) method was applied in the final runs due to superior comparison with the validation data compared to the Wigmosta *et al.* (1994) approach.

Table S 7 Comparison of Ding et al. (2014) (D) and Wigmosta et al. (1994) (W) precipitation partitions on percentage snowfall of total precipitation (Pr sno), mean daily snowmelt (S melt), mean daily ice melt (I melt) and mean daily ablation as modelled by T&C. Note that the Ding et al. (2014) precipitation partition was always used to apply the snowfall undercatch correction for Shallap and Artesonraju Glaciers so the quantity of input precipitation is the same in all runs. The Wigmosta et al. (1994) thresholds applied were T_{min} = -0.8 °C and T_{max} 2.9 °C.

Site	**Elevation**	**Pr sno**	**Pr sno**	**Pr sno**	**S melt mean**	**S melt mean**	**I melt mean**	**I melt mean**	**Ablation mean**	**Ablation mean**
Unit	m a.s.l.	%	%	%	mm w.e. day^{-1}	%	mm w.e. day^{-1}	%	m w.e. day^{-1}	%
Approach		D	W	D-W	D-W	D-W	D-W	D-W	D-W	D-W
SG	4790	94.4	54.4	40.0	2.0	42.4	-2.6	-19.1	-0.6	-3.4
AG	4797	95.0	43.3	51.7	3.1	54.6	-6.0	-39.0	-2.9	-13.7
CG	4821	91.3	67.5	23.8	1.0	25.5	-2.2	-20.7	-1.2	-8.0
QQG	5180	99.5	65.3	34.1	0.9	34.7	-1.4	-23.7	-0.6	-6.2
QIC	5650	99.8	96.5	3.3	0.0	-0.3	0.0	NaN	0.0	-1.0

Text S3.8 Snow and ice mass balance

The calculation of the evolution of the snow and ice mass balance is rather similar, so they will be treated together here. The calculations are performed for snow if there is snow precipitation during a timestep or the modelled snow water equivalent at the surface is greater than zero. Net input of energy to the snow or ice pack will increase its temperature, and after the temperature has been raised to the melting point, additional energy inputs will

result in melt. The change in the average temperature of the ice or snowpack (dT) is calculated using:

$$dT = \frac{1000\ dQ\ dt}{c_i \rho_w WE_b},$$

Where dt is the time step (h) and WE_b (mm w.e.) is the water equivalent mass of the ice or snowpack before melting. Energy inputs into an isothermal ice/snow pack result in melt (M in mm w.e.) calculated from:

$$M = \frac{1000\ dQ}{\lambda_f \rho_w}.$$

The water equivalent mass of the snow/ice pack after melting ($WE(t)$) is updated conserving the mass balance following:

$$WE(t) = WE(t - dt) + P_{r,sno}(t) - E(t)dt - M(t),$$

Here $E = \lambda E / \lambda_s$, which is the evaporation/sublimation from ice/snow in mm w.e. The snow density is assumed to be constant with depth and calculations are performed in a single snowpack layer. The snow density evolves over time using the method proposed by Verseghy (1991) and improved by Belair *et al.* (2003). In this parameterisation the snow density increases exponentially over time due to gravitational settling and is updated when fresh snow is added to the snowpack. Two parameters are required in this scheme, ρ_{sno}^{M1} and ρ_{sno}^{M2}, which represent the maximum snow density (kg m^{-3}) under melting and freezing conditions, respectively. These parameters typically have values of ρ_{sno}^{M1} = 500-600 kg m^{-3} and ρ_{sno}^{M2} = 300-450 kg m^{-3}. The depth of the ice pack can be increased through the formation of ice from the snowpack, which is prescribed to occur if the snow density increases to greater than ρ_{iceT} = 500 kg m^{-3} (a density associated with the firn to ice transition) and at a rate of 0.037 mm h^{-1} (Cuffey and Paterson, 2010). The density of ice is assumed constant with depth and given a value of 916.2 kg m^{-3}.

Text S3.9 Cuchillacocha albedo and incoming longwave radiation

For Cuchillacocha Glacier the snow albedo and incoming longwave radiation are calculated using the parameterisations described above using the standard parameters for the entire time series. To investigate their suitability the outputs were compared with just over 14 days of on-glacier measurements of hourly albedo (derived from incoming and outgoing shortwave radiation) and net radiation. There were only a few small snowfalls in the record, but nevertheless the comparison is good (Figure S 4a), with the NSE and RMSE between measured hourly albedo and modelled albedo equating to 0.60 and 0.10, respectively. The comparison with net radiation (Figure S 4b) includes the effect of both the albedo and incoming longwave parameterisation, with the NSE = 0.83 and RMSE = 76.4 W m^{-2}. These comparisons demonstrate that both the albedo parameterisation and incoming longwave parameterisation using the initial parameters are suitable for Cuchillacocha Glacier.

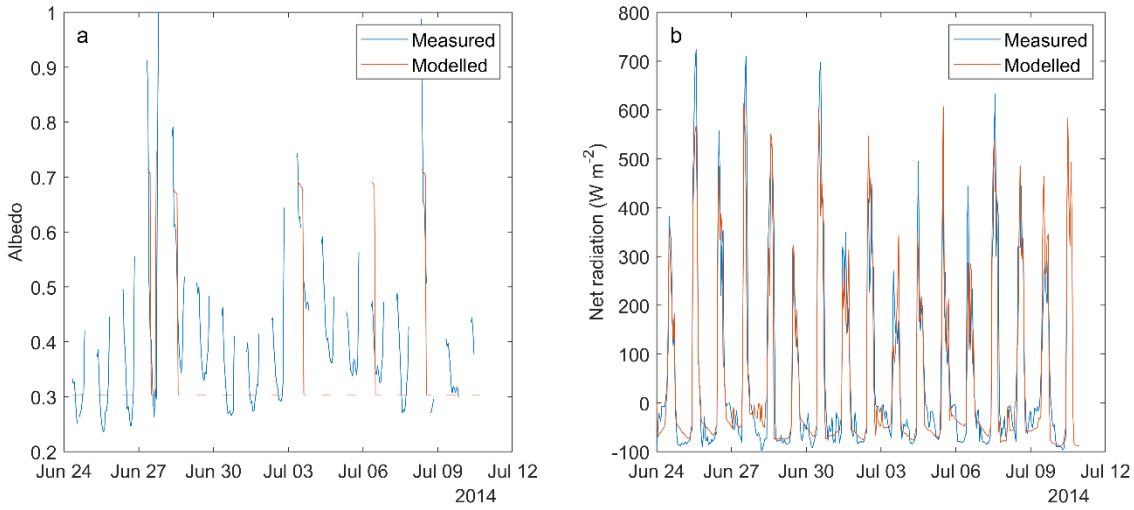

Figure S 4 Comparison of measured and T&C modelled albedo (panel a) and net radiation (panel b) at Cuchillacocha Glacier.

Table S 8 Parameters and their ranges included in the T&C Monte-Carlo simulations. *Different initial values were given for Quelccaya (first value) and the other sites (second value).

Parameter	**Range (initial)**	**Notes**	**References**
z_{0m}	0.001 – 0.05 (0.001) m	$z_{0h} = z_{0w} = 0.1 z_{0m}$	Covers range of z_{0h} and z_{0w} mentioned in Winkler et al. (2009) (without penitentes), Favier et al. (2004a), Sicart et al. (2011) and Gurgiser et al. (2013).
ε	0.95-0.99 (0.97)	Same value applied for snow and ice.	Since the natural variability of emissivity is small, perturbed by ±0.02.
Brock a	0.4-0.9 (0.713)	New snow albedo in deep snow equation.	Range following Pellicciotti (2004)
Brock b	0.1-0.17 (0.112)	Log multiplier in deep snow equation.	Range following Pellicciotti (2004)
Pr_t	3-15 (5/10*) mm	Threshold precipitation	Covers reasonable range either side of initial values.
Dilley and O'Brian a	30-60 (59.38) m^2 W^{-1}		Ranges span the range of optimal values found through calibration for different seasons in Juszak and Pellicciotti (2013).
Dilley and O'Brian b	110-170 (113.7) m^2 W^{-1} K^{-6}		
Dilley and O'Brian c	90-120 (96.96) m^3 $kg^{-0.5}$ W^{-1}		
ρ_{sno}^{M1}	450-650 (510/600*) kg m^{-3}	For QIC $\rho_{sno}^{M2} = \rho_{sno}^{M1}$, for other sites $\rho_{sno}^{M2} = \rho_{sno}^{M1} - 150$. $\rho_{iceT} = \rho_{sno}^{M1} - 5$.	Covers reasonable range either side of initial values.

Text S4. Results

Text S4.1 Model validation

Text S4.1.1 Comparison of SR50 records

SR50 (snow depth) data were cleaned to remove clearly erroneous data, with short spikes removed using a standard deviation filter and then filled using a running average either side

of the missing data. At Shallap and Artesonraju tilt of the SR50 relative to the surface is the likely cause of the erroneous data. This resulted in longer data gaps and may also cause some of the deviation with the modelled surface height in the retained data. However, at Quelccaya this is very unlikely as the station poles sit on steel plates set in ice/dense firn. Due to data gaps in the SR50 records and resetting of the SR50 during site visits the records are not continuous, for this reason the values of surface height were set equal to the model values at the beginning of each time period.

Text S4.1.2 Using albedo as validation

Where measured albedo derived from on-glacier shortwave measurements is available this can be used to estimate the surface type (ice or snow) and therefore be compared with the modelled surface. Measured albedo for the purposes of this comparison is calculated as the accumulated albedo (or the 24 hour sum of $S\uparrow$ divided by the sum of $S\downarrow$, centred over the time of observation (van den Broeke et al., 2004)), since this reduces the effect of the sun angle on albedo observations. A simple albedo threshold of 0.4 is used to determine if the surface is likely ice (below) or snow (above). This is based on the observations within Brock et al. (2000) that ice albedo is generally less than 0.3, and snow albedo is usually greater than ~0.5. The 'measured' surface type is compared to the modelled surface on both an hourly basis and daily. For the daily comparison the measured surface is that at 12:00 and the modelled surface is given as snow if more than 12 hours of the day have a snow surface. To assess the model performance over the full record at Cuchillacocha all useable visible imagery of the glacier during the modelling period was downloaded from Planet Labs (20 images in total), and the surface type at the meteorological station (so whether ice or snow) was identified by eye from the imagery.

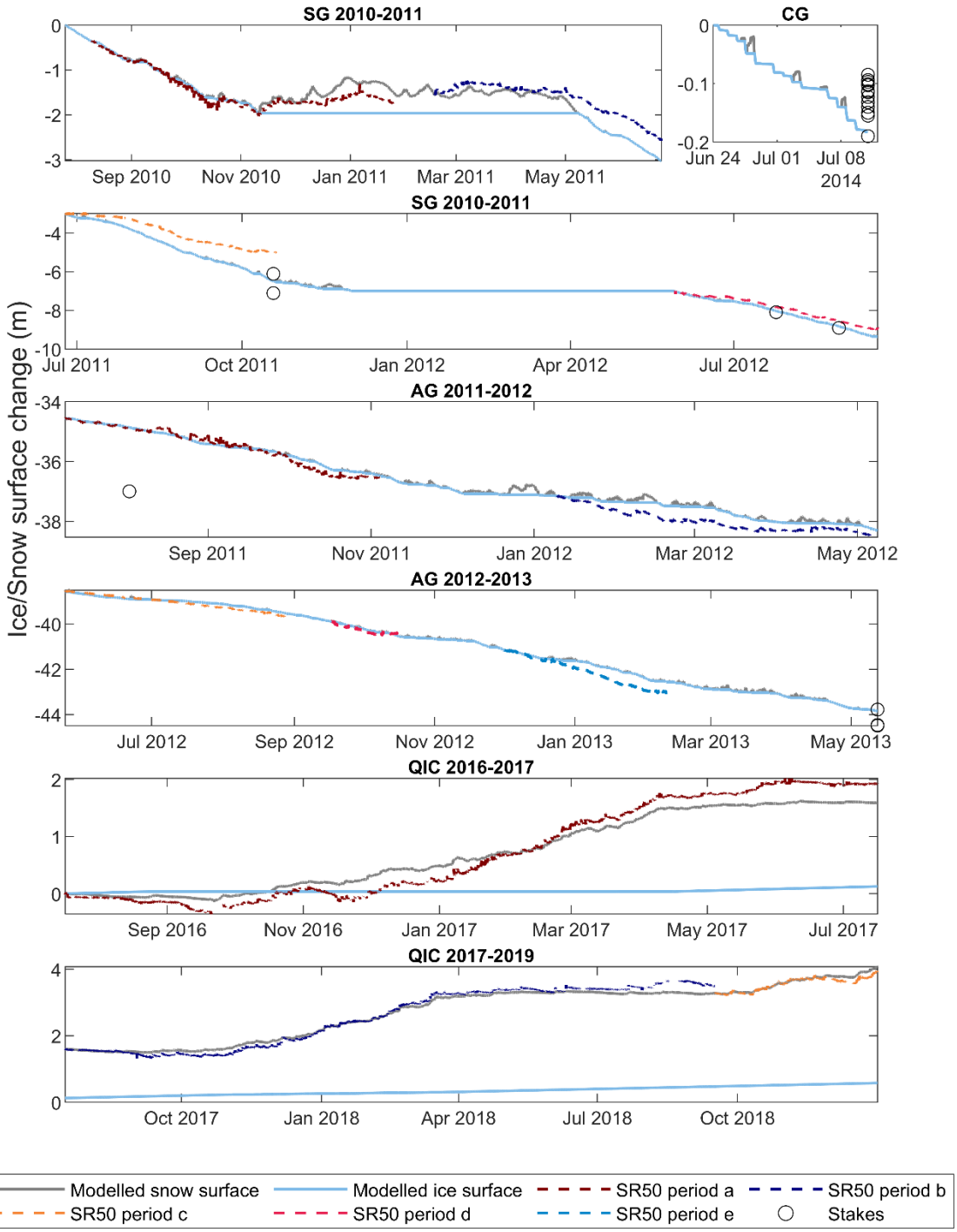

Figure S 5 Comparison of T&C modelled ice/snow surface against validation data from SR50 and stake measurements. Note that the Artesonraju and Cuchillacocha modelling periods are longer than shown, only the period with validation data is represented. To allow SR50 and stake data to be represented together surface change is in actual m not m of water equivalent.

Table S 9 SR50 validation data. The difference between the SR50 data and the T&C modelled ice/snow surface (Diff.) is calculated as modelled minus measured so that positive difference values demonstrate that the model is overestimating the surface height. The SR50 periods are shown in Figure S 5 and values are not in water equivalent.

SR50 validation			
SR50 period	**ME (m)**	**RMSE (m)**	**Diff. (m)**
SG			
a	0.114	0.167	0.264
b	-0.245	0.294	-0.477
c	-0.672	0.757	-1.542
d	-0.200	0.223	-0.456
AG			
a	0.041	0.136	0.110
b	0.359	0.387	0.232
c	0.053	0.104	0.142
d	0.070	0.088	-0.106
e	0.322	0.384	0.483
QIC			
a	-0.022	0.202	-0.330
b	-0.022	0.130	-0.219
c	0.057	0.092	0.140

Table S 10 Stake validation data. The difference is calculated as measured minus T&C modelled ablation (with ablation a negative value), so that positive values represent an overestimation of ablation by the model. The Cuchillacocha stakes were all within 10 m of the weather station, so they are given the weather station coordinates, similarly the STAT stakes on Shallap were adjacent to the weather station, with SH-7 and SH-8 being the closest two stakes to the station in the 2010-2011 season (being 66 m and 148 m from the station, respectively). For Artesonraju the stakes surrounding the station are used for validation (being 27-87 m from the station), of which only A14 was measured in 2011. *The stakes were measured after the end of the modelling period (22/08/2013) so the values were adjusted pro rata to the end of the modelling period.

Stake validation						
Stake	**Lat. (°)**	**Long. (°)**	**Start Date**	**End Date**	**Diff. (m)**	**Diff. (m w.e.)**
SG						
SH-7	-9.4888	-77.3376	22/07/2011	18/10/2011	-0.615	-0.564
SH-8	-9.4887	-77.3367	22/07/2011	18/10/2011	0.382	0.350
STAT	-9.4892	-77.3380	20/06/2012	24/07/2012	-0.074	-0.068
STAT	-9.4892	-77.3380	24/07/2012	28/08/2012	-0.093	-0.085
AG						
A14	-8.9655	-77.6356	28/04/2011	02/08/2011	-2.141	-1.962
A13	-8.9642	-77.6357	27/09/2012	12/05/2013*	0.067	0.061
A14	-8.9650	-77.6358	27/09/2012	12/05/2013*	-0.644	-0.590
A18	-8.9651	-77.6350	27/09/2012	12/05/2013*	-0.617	-0.565
CG						
CG 1	-9.4054	-77.3521	24/06/2014	10/07/2014	0.043	0.040
CG 2	-9.4054	-77.3521	24/06/2014	10/07/2014	0.033	0.030
CG 3	-9.4054	-77.3521	24/06/2014	10/07/2014	-0.007	-0.006
CG 4	-9.4054	-77.3521	24/06/2014	10/07/2014	0.086	0.079
CG 5	-9.4054	-77.3521	24/06/2014	10/07/2014	0.082	0.075
CG 6	-9.4054	-77.3521	24/06/2014	10/07/2014	0.089	0.082
CG 7	-9.4054	-77.3521	24/06/2014	10/07/2014	0.070	0.064
CG 8	-9.4054	-77.3521	24/06/2014	10/07/2014	0.098	0.090
CG 9	-9.4054	-77.3521	24/06/2014	10/07/2014	0.068	0.062
CG 10	-9.4054	-77.3521	24/06/2014	10/07/2014	0.080	0.073
CG 11	-9.4054	-77.3521	24/06/2014	10/07/2014	0.028	0.026
CG 12	-9.4054	-77.3521	24/06/2014	10/07/2014	0.051	0.047
CG Mean	-9.4054	-77.3521	24/06/2014	10/07/2014	0.060	0.055

Table S 11 Surface type validation. The 'Corr.' Columns represent when the T&C model correctly replicated the surface type and the 'In.' columns represent when the model incorrectly replicated the surface type.

Surface type validation				
Hourly	Corr. ice (%)	Corr. snow (%)	In. ice (%)	In. snow (%)
SG	42.3	42.2	8.8	6.7
AG	40.7	25.0	11.6	22.7
CG	71.7	4.9	10.9	12.5
CG (PL)	52.6	15.8	26.3	5.3
QQG	15.4	38.9	39.4	6.4
Daily	Corr. ice (%)	Corr. snow (%)	In. ice (%)	In. snow (%)
SG	45.8	42.2	5.6	6.4
AG	42.4	25.2	11.6	20.8
CG	76.9	7.7	7.7	7.7
QQG	16.1	37.8	40.5	5.7

Text S4.2 Monte-Carlo analysis

The results of the Monte-Carlo analysis can be seen in Figure S 6. The uncertainty (based on not water equivalent comparisons) given by the range in the Monte-Carlo T&C runs for Shallap (16.3%), Artesonraju (15.7%) and Quisoquipina (18.8%) is likely due to variations in the surface roughness, since at these sites the run with the initial values was close to the minimum Monte-Carlo run and the initial surface roughness was the minimum for the range. At Cuchillacocha and Quelccaya the uncertainty was much larger (129.6% and 245.3%, respectively). This is because both albedo and incoming longwave were modelled at these sites, and at Quelccaya the thick snowpack and low ablation rates mean results are very sensitive to the snow density. However, comparison of available measured albedo and incoming longwave radiation at Cuchillacocha showed very close agreement with modelled values (see Text S3.9), and the calibration at Quelccaya was used to determine the best Prt and snow density parameter values. We note that the ranges we used are large, resulting in rather large (conservative) uncertainties. This analysis demonstrates that there is most confidence in the results from Shallap, Artesonraju and Quisoquipina, and shows why checks with additional data (Cuchillacocha) and calibration (Quelccaya) were necessary.

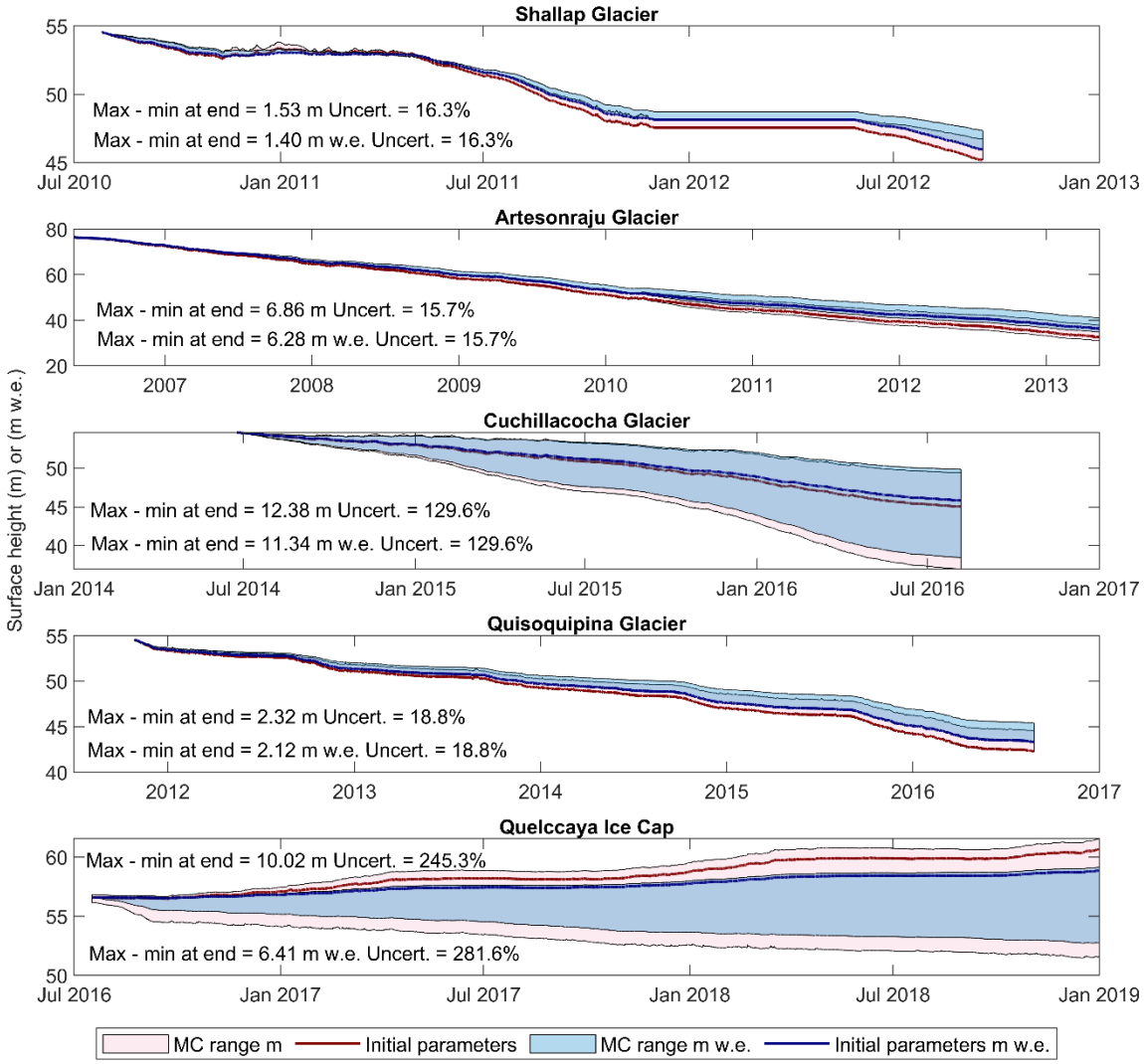

Figure S 6 Monte-Carlo parameter uncertainty. The 'MC range' relates to the range of T&C modelled surface height values across all 1000 Monte Carlo runs, given both in metres and metres water equivalent. The run using the initial parameters (those used throughout the rest of the analysis) is also shown, alongside the difference (in m and in m w.e. and as a percentage of the total mass loss using the initial parameters) between the maximum and minimum surface height at the end of the study period.

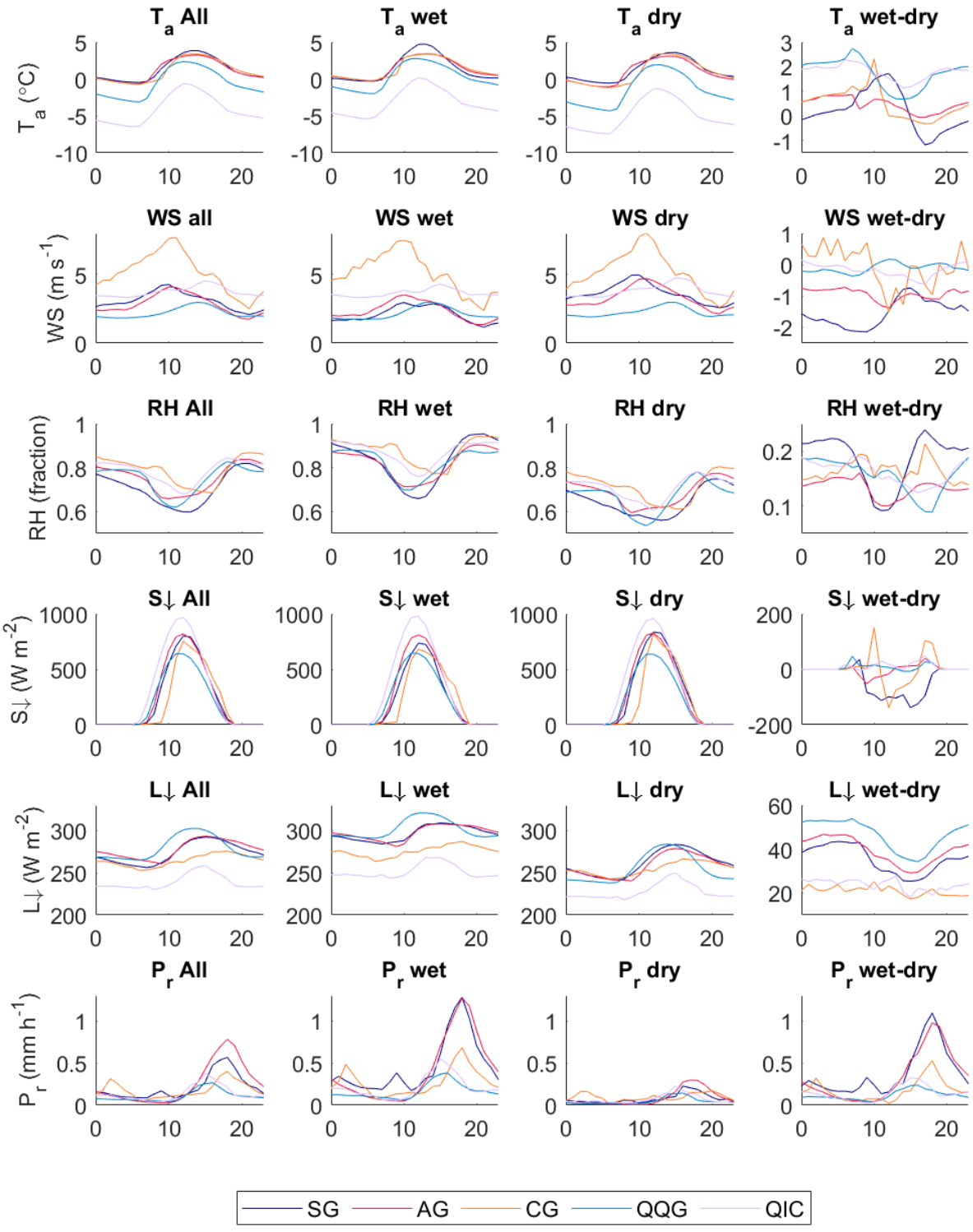

Figure S 7 Diurnal and seasonal variation in the input meteorological variables at each site. Most of the data were measured at the on-glacier stations, except that filled or modelled as explained in Text S2.3. Note that the small precipitation peak at Shallap (SG) at 09:00 is likely due to melt of snow that has stuck to the rim of the gauge and the early morning peak in precipitation at Cuchillacocha (CG) is because this is when snow fall was 'moved' to which otherwise collected in the unheated gauge and melted mid-morning (see Text S2.3). T_a is air temperature, WS is wind speed, RH is relative humidity, $S\downarrow$ is incoming shortwave radiation, $L\downarrow$ is incoming longwave radiation and P_r is precipitation.

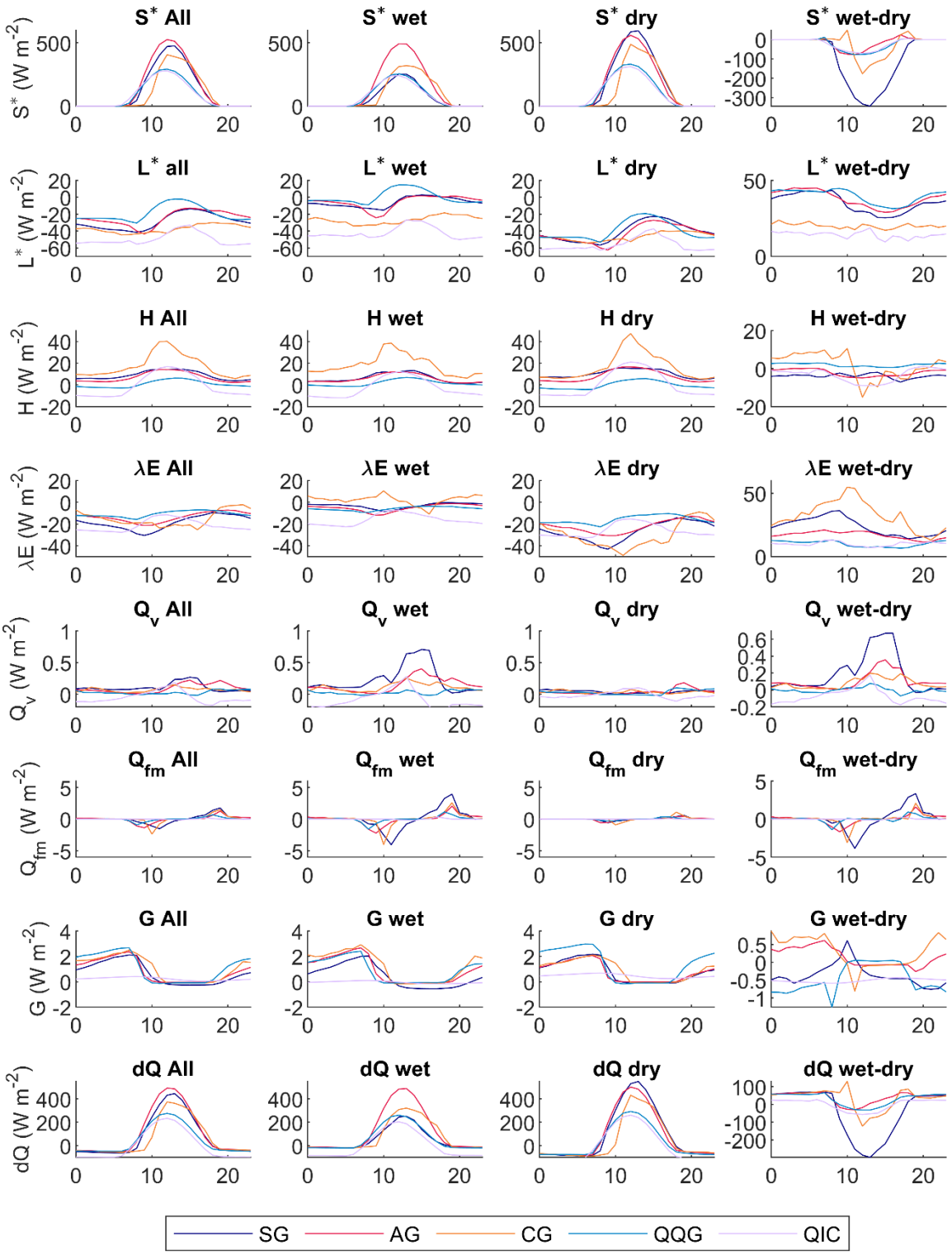

Figure S 8 Diurnal cycle of average hourly fluxes for all glaciers derived from T&C outputs. Data has been averaged over the entire period in the left most column, and over the wet and dry season in the middle columns, with the right most column showing the difference between the wet and dry season. S* is net shortwave radiation, L* is net longwave radiation, H is the sensible heat flux, λE is the latent heat flux, Qv is the heat flux due to precipitation, Qfm is the heat flux due to melting and freezing of water in the snowpack, G is the ground heat flux and dQ is the energy available for melt.

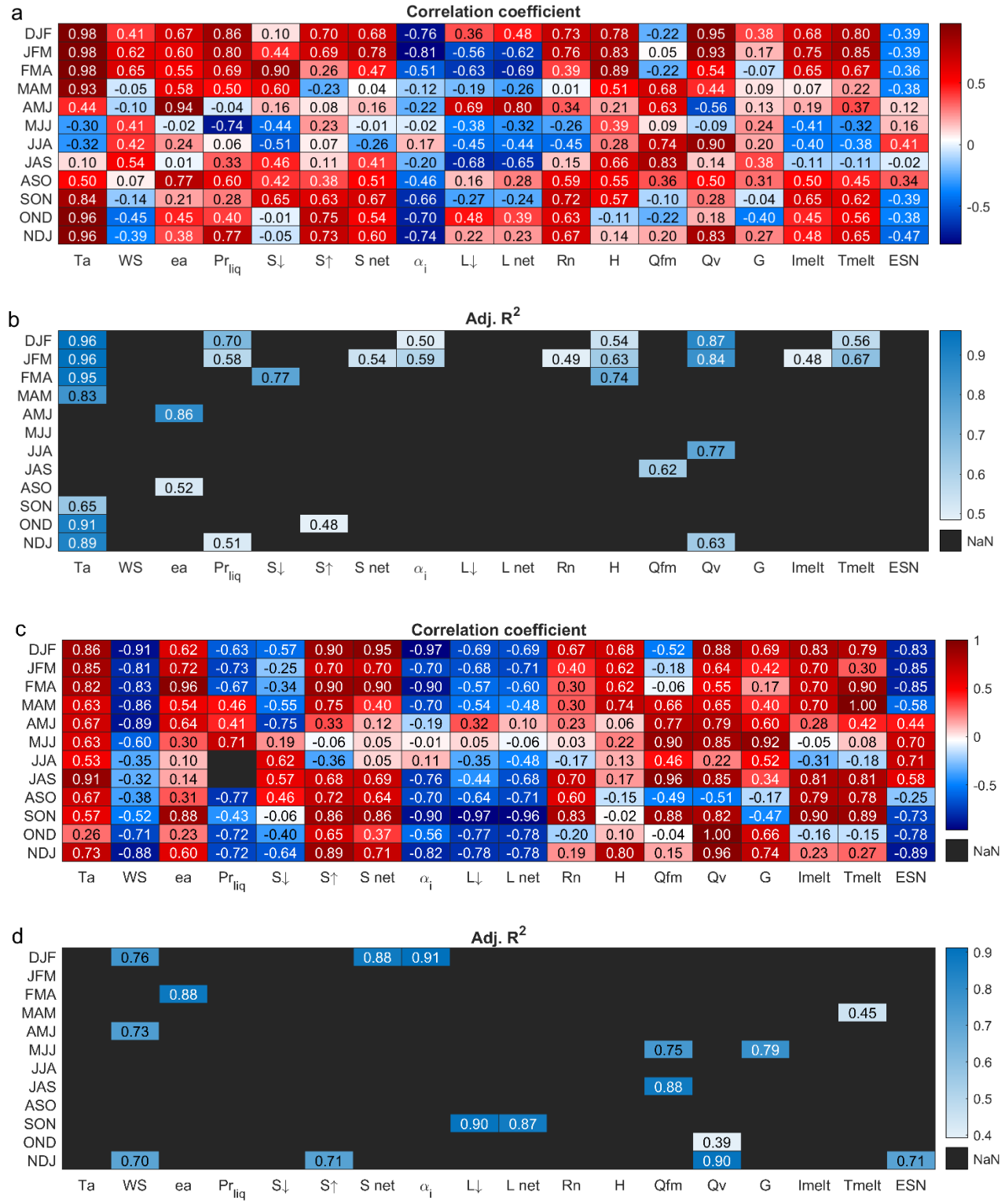

Figure S 9 Statistical analysis of the relationship between meteorological, energy flux and mass balance variables from T&C and the NOAA ONI. Here, Ta is air temperature, WS is wind speed, Pr_{liq} is rainfall, $S\downarrow$ is incoming shortwave radiation, $S\uparrow$ is outgoing shortwave radiation, S net is net shortwave radiation, α_i is albedo, $L\downarrow$ is incoming longwave radiation, L net is net longwave radiation, Rn is net radiation, H is the sensible heat flux, Qfm is the heat flux from melting or freezing of water in snow, Qv is the heat flux due to precipitation, G is the ground heat flux, Imelt is ice melt, Tmelt is total melt and ESN is sublimation from snow. Panels a) and b) are for Artesonraju and panels c) and d) are for Quisoquipina. The correlation coefficient given (panels a) and c)) was

derived from a Pearsons correlation if the data were normally distributed and from a Spearman's rank correlation if the data were not normally distributed. The adjusted R^2 (panels b) and d)) is from the linear regression. Regression was only conducted when the correlation was significant at $p<0.05$. The statistical results for several variables which did not give any significant relationships at any site have not been shown. These include relative humidity, total precipitation, solid precipitation, outgoing longwave radiation, latent heat flux, snow melt and sublimation from ice.

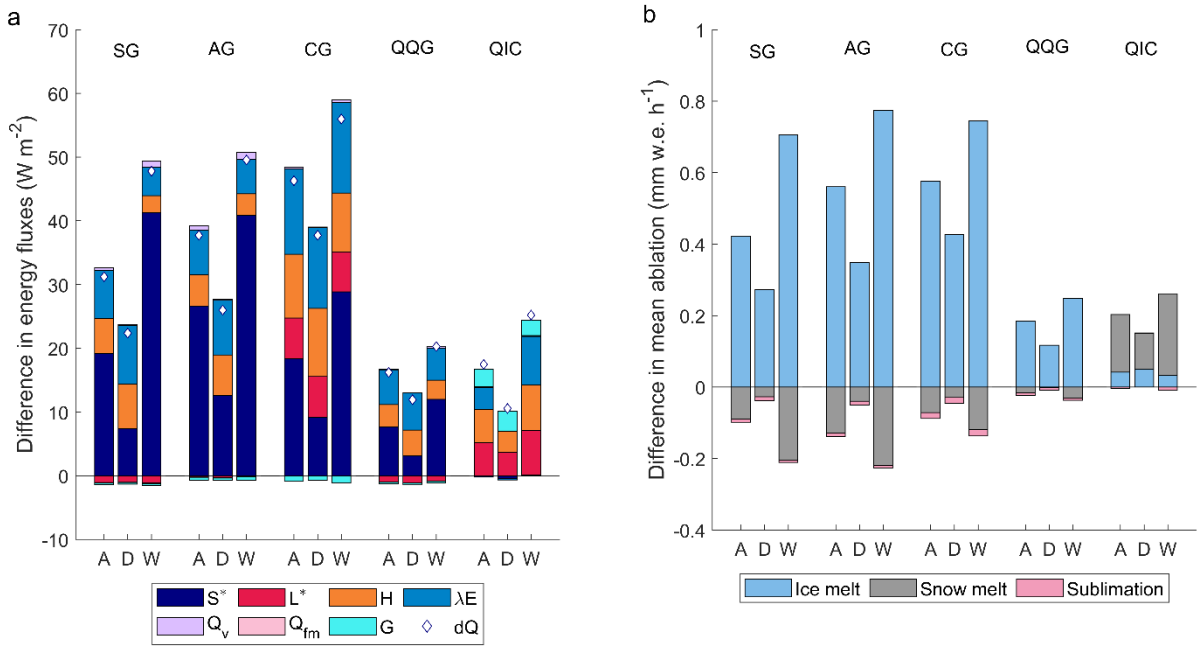

Figure S 10 Difference between T&C standard run and an increase in air temperature of 2°C (scenario minus standard run). Panel a) shows the difference in the mean total and seasonal energy balance fluxes and b) shows the difference in the mean total and seasonal ice melt, snow melt and sublimation. S* is net shortwave radiation, L* is net longwave radiation, H is the sensible heat flux, λE is the latent heat flux, Qv is heat flux due to precipitation, Qfm is the heat flux due to melting and freezing of water in the snowpack, G is the ground heat flux and dQ is the energy available for melt. 'A' is the difference averaged over the whole record, 'D' is the difference for the dry season, and 'W' is the difference for the wet season.

Text S5. Discussion

Table S 12 Comparison of meteorological, energy balance and ablation conditions of South American glaciers, ordered by latitude. The values are means over the entire record. For the Peruvian sites albedo is the hourly albedo used in the model (see text for its derivation). The Chilean glaciers' albedo is average hourly albedo, with their fluxes from the 'Reference database' (see Schaefer et al. (2020)). Ele. is the station elevation. The class is the climate group as defined in Sagredo and Lowell (2012). Ta is air temperature, RH is relative humidity, WS is wind speed, α is albedo, S* is net shortwave radiation, L* is net longwave radiation, H is sensible heat, λE is the latent heat flux and Sub is sublimation.

Glacier	**Ref.**	**Lat.**	**Long.**	**Country**	**Ele.**	**Time period**	**Class**	**Ta**	**RH**	**WS**	**α**	**S***	**L***	**H**	**λE**	**Melt**	**Sub**
		(°)			(m a.s.l)	(dd/mm/yyyy)		(°C)	(%)	(ms^{-1})		($W\ m^{-2}$)				(mm w.e. h^{-1})	
Antizana 15	Favier et al. (2004b)	-0.47	-78.15	Ecuador	4890	14/03/02-14/03/2003	1	0.3	81	4.8	0.49	123	-39	21	-27	0.845	0.034
Artesonraju	This paper	-8.96	-77.64	Peru	4797	20/05/06–12/05/2013	2.1	1.1	75	2.8	0.44	141	-25	7	-14	0.877	0.017
Cuchillacocha	This paper	-9.41	-77.35	Peru	4821	24/06/14-05/08/2018	2.1	1.1	79	5.1	0.49	100	-37	17	-15	0.616	0.019
Shallap	This paper	-9.49	-77.34	Peru	4790	26/07/2010 – 31/11/2012 and 29/05/2012 – 18/09/2012	2.1	1.7	71	3.1	0.52	123	-29	9	-19	0.765	0.024
Quisoquipina	This paper	-13.79	-70.89	Peru	5180	21/10/11–25/08/2016	2.2	-0.6	75	2.2	0.61	81	-19	1	-11	0.362	0.014
Quelccaya	This paper	-13.92	-70.82	Peru	5650	17/07/16–31/12/2018	2.2	-4.1	78	3.7	0.71	80	-50	-1	-21	0.006	0.027
Zongo	Favier et al. (2004b); Sicart et al. (2002)	-16.25	-68.17	Bolivia	5050	1/09/99-31/08/2000	2.2	-0.8	71	2.7	0.66	72	-45	21	-31	0.205	0.039
Guanaco 08/09	Ayala et al. (2017)	-29.349	-70.018	Chile	5324	Dec 2008 to Jan 2009	4	-5.3	43	5.9	-	192	-104	25	-70	0.008	0.071
Guanaco 09/10	Ayala et al. (2017)	-29.349	-70.018	Chile	5324	Dec 2009 to Jan 2010	4	-5.0	45	7.3	-	189	-98	24	-90	0.033	0.083
Guanaco 10/11	Ayala et al. (2017)	-29.349	-70.018	Chile	5324	Dec 2010 to Jan 2011	4	-6.3	39	6.9	-	228	-112	21	-84	0.000	0.092
Tapado	Ayala et al. (2017)	-30.148	-69.925	Chile	4775	Dec 2013 - Jan 2014	4	-	-	-	-	233	-82	63	-115	1.296	0.067
Juncal Norte 3305 m	Ayala et al. (2017)	-32.982	-70.114	Chile	3305	Dec 2008 - Jan 2009	5	5.7	51	3.3	-	255	-80	51	-19	2.438	0.008
Juncal Norte 3127 m	Ayala et al. (2017)	-32.991	-70.109	Chile	3127	Dec 2008 - Jan 2009	5	-	-	-	-	279	-78	66	-19	2.800	0.008

Yeso 13/14	Ayala et al. (2017)	-33.529	-69.92	Chile	4428	Dec 2013 - Jan 2014	5	0.0	48	2.0	-	214	-67	37	-65	1.379	0.042
Yeso 14/15	Ayala et al. (2017)	-33.529	-69.92	Chile	4428	Dec 2014 - Jan 2015	5	-	-	-	-	172	-75	12	-37	1.121	0.021
Bello	Schaefer et al. (2020)	-33.53	-69.94	Chile	4134	01/01/15-31/03/2015	5	2.3	37	2.9	0.26	223	-69	25	-22	1.681	-
San	Schaefer et al. (2020)	-33.75	-70.07	Chile	3466	01/03/16-31/03/2016	5	7.1	43	2	0.37	137	-42	11	-2	1.108	-
Mocho	Schaefer et al. (2020)	-39.94	-72.02	Chile	2003	31/01/06-21/03/2006	6.1	5.9	66	6.3	0.57	118	-11	59	-4	1.758	-
Exploradores	Schaefer et al. (2020)	-46.51	-73.18	Chile	191	01/01/15-01/03/2015	6.2	7.4	87	3.1	0.23	143	-2	65	38	2.636	-
Tyndall 2015	Schaefer et al. (2020)	-51.13	-73.31	Chile	608	01/01/15-01/03/2015	7	4.8	74	5.6	0.23	94	-14	65	9	1.719	-
Tyndall 2016	Schaefer et al. (2020)	-51.13	-73.31	Chile	608	01/01/16-01/03/2016	7	5.3	72	5.7	0.43	110	-13	76	9	1.949	-