Supplementary Information

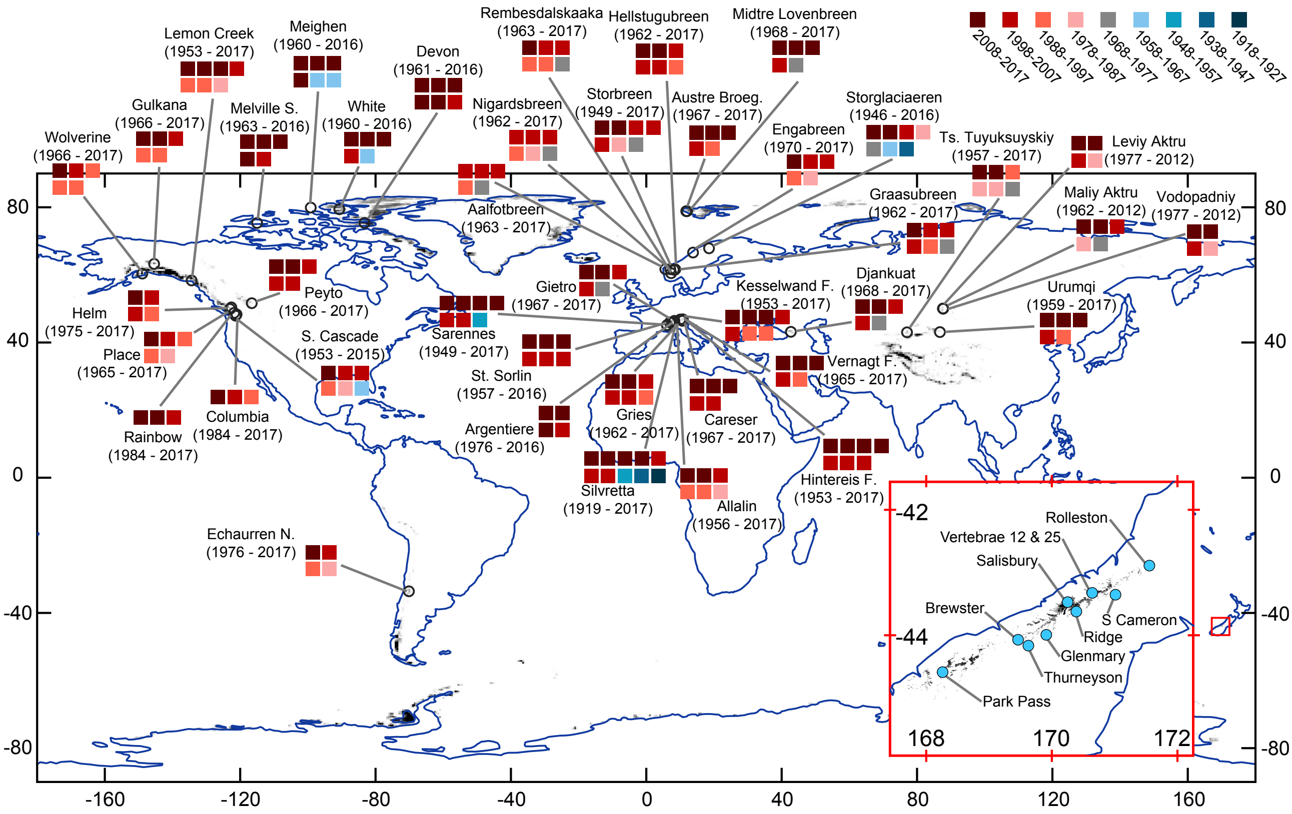

Figure S.1. **Global glaciers, increasing extreme mass-loss years in recent decades, and study area. Full map: All global glaciers[53] (black), including all glaciers with >30 years of mass-balance measurements[11] (open black circles). Each colored square represents one extreme mass-loss year, with the color showing the timing, by decade. Extreme mass-loss years are defined as 90th percentile of negative mass balances for each individual glacier. Inset: Glaciers in the Southern Alps of New Zealand[53] (black), including the subset of glaciers used in this study (blue).**

Model calibration. Brewster Glacier is the only glacier in this study with weather station data. Comparison of weather station data collected from below Brewster Glacier (June 2004 – November 2009)[41] with VCSN climate for the same period showed VCSN temperature should be reduced by 1.25°C, and precipitation should be increased by a factor of 1.3 to match the weather station data. We then used measured mass-balance and snowline data from Brewster Glacier (2005 – 2017)[24] to calculate optimum temperature (M_T) and radiation (M_R) factors.

We perform a grid search for optimum parameters, shown in Table S.1, chosen following previously published temperature and radiation factors that have been shown to be transferable across glaciers[54]. Optimum temperature (M_T) and radiation (M_R) factors of 1.46 mm d^{-1} °C^{-1} and 0.15 m^2 mm W^{-1} d^{-1}, respectively, result in an annual root mean square error (RMSE) of 560 mm w.e. and series mean error of 4 mm w.e. for mass balance. To estimate uncertainties in the model parameters, we use both the annual RMSE and series mean error to select best parameter combinations, and a threshold of +50% of the minimum annual RMSE for each. For Brewster Glacier, we use all parameter combinations where the annual RMSE is less than 840

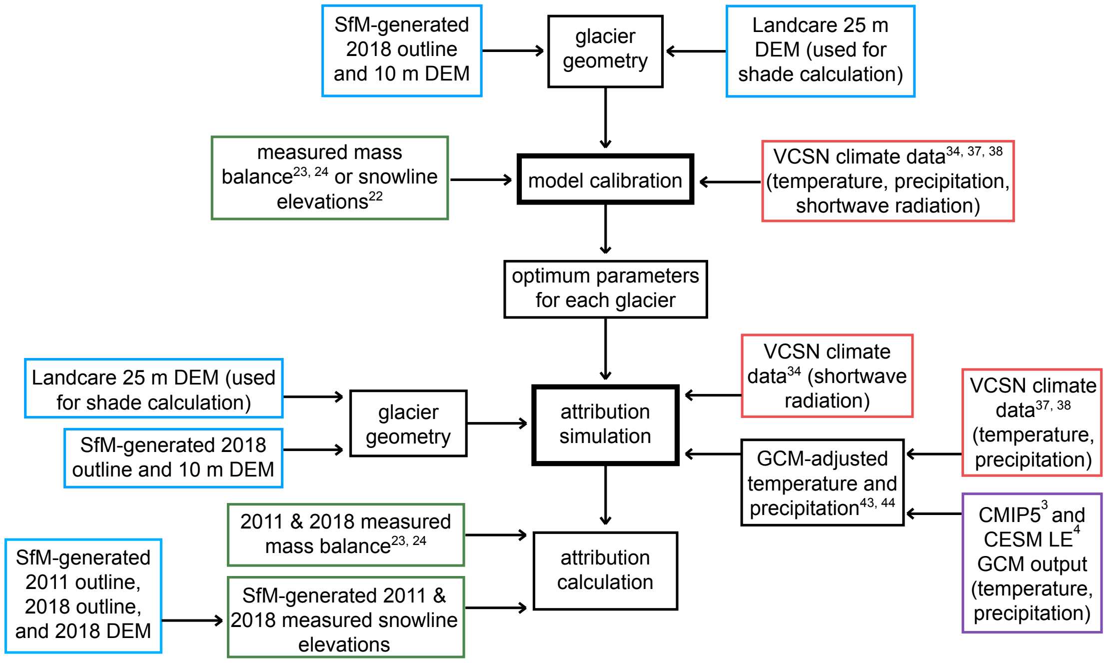

Figure S.2. Methodology flow chart. Flow chart showing the input data and steps in the methodology, to ultimately calculate attributions for mass balance and snowlines. Colors indicate different types of data (green: measured mass balance and snowlines; blue: glacier geometry; red: climate data; purple: GCM output). Structure from motion photogrammetry is abbreviated to SfM.

(560+280) mm w.e. and series mean error is less than 280 mm w.e., resulting in a suite of 46 parameter combinations for which attributions are calculated (Table S.2, Figs. S.4a & S.5a).

We calibrate snowlines separately from mass balance for Brewster Glacier, but over the same time period (2005 – 2017) and for the same ranges of M_T and M_R for comparable calibrations (Table S.1). For snowlines, again we use both the annual RMSE and series mean error to select best parameter combinations, and a threshold of +50% of the minimum annual RMSE for each. For Brewster Glacier, we use all parameter where the annual RMSE is less than 117 (78+39) m and series mean error is less than 39 m, resulting in a suite of 34 parameter combinations for which attributions are calculated (Table S.2, Figs. S.4b & S.5b). Despite the separate calibrations between snowlines and mass balance, of the 34 best-fitting parameter combinations for the snowline calibrations, 29 of those parameter combinations are also best-fitting parameter combinations for mass balance.

Climate adjustments and model parameters at Brewster Glacier are not consistent for all New Zealand glaciers. For all other glaciers without weather station data we add two additional parameters: an additive temperature adjustment (T_{adj}; °C) and multiplicative precipitation adjustment (P_{adj}; multiplicative), which are used to adjust VCSN temperature and precipitation data, respectively. For additional glaciers, the calibration setup (Table S.1) and calibrations results (Table S.2, Figs. S.4 & S.5) are shown.

The results of the calibration, including the annual RMSE and series mean error ranges used for the suite of best-fitting parameters, and ranges of parameters, vary among the glaciers (Table S.2). Glacier size influences annual snowline RMSE,

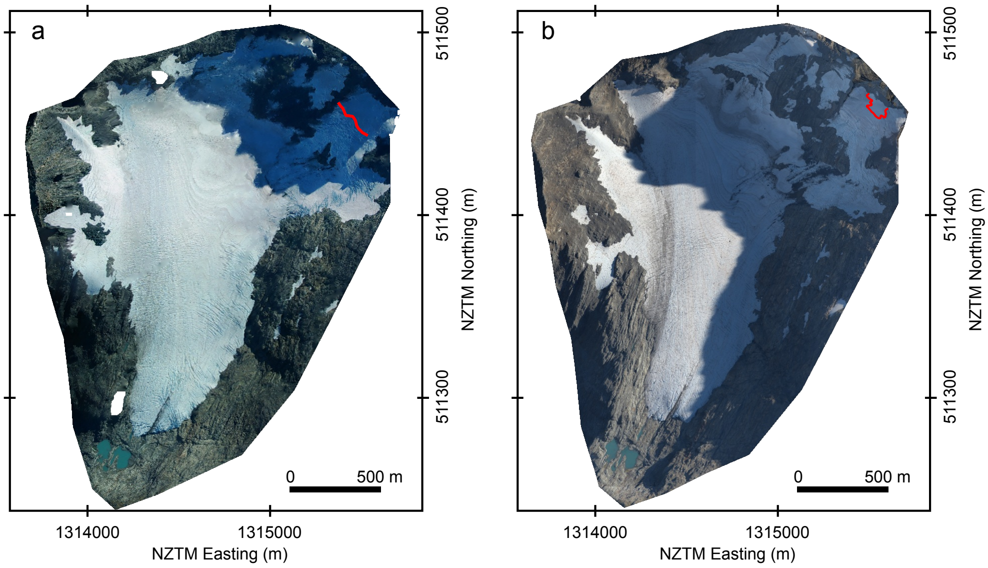

Figure S.3. Example structure from motion photogrammetry orthophotos. Brewster Glacier orthophotos used to measure snowlines, with digitized snowlines shown in red. (a) 2011 orthophoto mosaic generated using four images taken in March 2011 and georeferenced with 2018 images, and (b) 2018 orthophoto mosaic generated and georeferenced using 192 geotagged images taken in March 2018.

as glaciers with smaller elevation range have a limited range of errors. The annual snowline RMSE for Rolleston Glacier for the suite of parameters ranges from 35 – 47 m (Table S.2). However, comparison of the measured and modeled snowlines shows that some years modeled snowlines are far from measured elevations (Fig. S.4d), and larger variation from the 1:1 line compared with other glaciers (Fig. S.5d). Comparatively, South Cameron Glacier annual snowline RMSEs for the suite of parameters ranges from 88 – 114 m (Table S.2), about double Rolleston Glacier. However, annual comparison of measured and modeled snowlines shows a better fit (Fig. S.4g), with points falling closer to the 1:1 line (Fig. S.5g).

Table S.1. Calibration setup for each glacier, including the full range tested for each parameter (step sizes for ranges shown in parenthesis).

Glacier -	Calib. period	M_T range (mm d^{-1} °C^{-1})	M_R range (m^2 mm W^{-1} d^{-1})	T_{adj} range (°C)	P_{adj} range (x)
Brewster MB	2005 – 2017	0.5 – 1.7 (0.08)	0.13 – 0.25 (0.02)	-1.25	1.3
Brewster SL	2005 – 2017	0.5 – 1.7 (0.08)	0.13 – 0.25 (0.02)	-1.25	1.3
Rolleston MB	2011 – 2016	0.5 – 1.7 (0.3)	0.13 – 0.25 (0.03)	-0.65 – -2 (0.45)	0.8 – 1.7 (0.45)
Rolleston SL	1981 – 2017	0.5 – 1.7 (0.3)	0.13 – 0.22 (0.03)	-0.65 – -2 (0.45)	0.8 – 1.25 (0.45)
Salisbury SL	1981 – 2017	0.5 – 1.7 (0.2)	0.13 – 0.21 (0.02)	-0.8 – -2 (0.4)	0.8 – 1.6 (0.4)
Park Pass SL	1981 – 2017	0.5 – 1.7 (0.4)	0.13 – 0.21 (0.04)	-0.8 – -2 (0.4)	0.8 – 1.6 (0.4)
South Cameron SL	1981 – 2017	0.5 – 1.7 (0.4)	0.13 – 0.21 (0.04)	-0.8 – -2 (0.4)	0.8 – 1.6 (0.4)
Thurneyson SL	1981 – 2017	0.5 – 1.7 (0.4)	0.13 – 0.21 (0.04)	-0.8 – -2 (0.4)	0.8 – 1.6 (0.4)
Vertebrae12 SL	1981 – 2017	0.5 – 1.7 (0.4)	0.13 – 0.21 (0.04)	-0.8 – -2 (0.4)	0.8 – 1.6 (0.4)
Ridge SL	1981 – 2017	0.5 – 1.7 (0.4)	0.13 – 0.21 (0.04)	-0.8 – -2 (0.4)	0.8 – 1.6 (0.4)
Glenmary SL	1981 – 2017	0.5 – 1.7 (0.4)	0.13 – 0.21 (0.04)	-0.8 – -2 (0.4)	0.8 – 1.6 (0.4)
Vertebrae25 SL	1981 – 2017	0.5 – 1.7 (0.4)	0.13 – 0.21 (0.04)	-0.8 – -2 (0.4)	0.8 – 1.6 (0.4)

Model validation and sensitivity. Following the calibration, we identify the parameter combination resulting in the lowest

Table S.2. Calibration results for each glacier, for parameter combinations resulting in lowest RMSEs and used in the attribution calculations. Includes minimum and maximum annual RMSE and series mean error for parameter combinations, and ranges of each parameter.

Glacier	Annual	Series	M_T	M_R	T_{adj}	P_{adj}
-	RMSE	mean error	(mm d^{-1} °C^{-1})	(m^2 mm W^{-1} d^{-1})	(°C)	(x)
Brewster MB	560 – 647 mm w.e.	4 – 280 mm w.e.	0.5 – 1.7	0.13 – 0.25	-1.25	1.3
Brewster SL	78 – 106 m	1 – 37 m	0.5 – 1.7	0.13 – 0.25	-1.25	1.3
Rolleston MB	583 – 746 mm w.e.	2 – 275 mm w.e.	0.5 – 1.7	0.13 – 0.25	-1.55 – -2	0.8 – 1.25
Rolleston SL	35 – 47 m	0 – 16 m	0.5 – 1.7	0.13 – 0.22	-0.65 – -1.55	0.8 – 1.25
Salisbury SL	106 – 124 m	2 – 49 m	0.9 – 1.7	0.13 – 0.21	-0.8 – -1.2	0.8
Park Pass SL	128 – 155 m	1 – 63 m	0.5 – 1.7	0.13 – 0.21	-0.8 – -2	0.8 – 1.2
South Cameron SL	88 – 114 m	4 – 40 m	0.5 – 1.7	0.13 – 0.21	-0.8 – -2	0.8 – 1.2
Thurneyson SL	79 – 100 m	1 – 33 m	0.5 – 1.7	0.13 – 0.21	-0.8 – -2	1.2 – 1.6
Vertebrae12 SL	112 – 121 m	10 – 52 m	0.9 – 1.7	0.13 – 0.21	-0.8 – -1.2	0.8
Ridge SL	87 – 109 m	3 – 42 m	0.5 – 1.7	0.13 – 0.21	-0.8 – -2	0.8 – 1.6
Glenmary SL	66 – 78 m	0 – 30 m	0.5 – 1.7	0.13 – 0.21	-0.8 – -2	1.2 – 1.8
Vertebrae25 SL	85 – 95 m	0 – 42 m	0.5 – 1.7	0.13 – 0.21	-0.8 – -1.6	0.8 – 1.2

annual RMSE and series mean error between modeled and measured mass balance or snowlines. To validate and evaluate the sensitivity of this calibration, we leave one of the years out, find the new optimum calibration for the model for the remaining years, and find the RMSE between the modeled and measured mass balance or snowline of the year left out. This is repeated for all years. We then calculated the variation (RMSE) between annual mass balance or snowlines calculated for optimum parameters, and annual mass balance or snowlines calculated when each year was left out. Results are shown in Table S.3. Validation results for glaciers with mass balance show a lower mean variation for Brewster Glacier (60 mm w.e.) compared with Rolleston Glacier (510 mm w.e.). This may be partially due to the shorter mass-balance record for Rolleston Glacier (6 years) compared with Brewster Glacier (13 years). Validation results for snowline records range from mean variations of 29 m (Rolleston Glacier) to 230 m (Thurneyson Glacier).

Table S.3. Model validation and sensitivity for each glacier, calculated by leaving each year out of the calibration (described in **Model validation** in the methodology).

Glacier	Calib.	Mean	Annual	Annual	Standard
-	period	variation	min.	max.	deviation
Brewster MB	2005 – 2017	60 mm w.e.	0 mm w.e.	144 mm w.e.	42 mm w.e.
Brewster SL	2005 – 2017	82 m	9 m	187 m	54 m
Rolleston SL	1981 – 2017	29 m	0 m	85 m	28 m
Rolleston MB	2011 – 2016	510 mm w.e.	11 mm w.e.	866 mm w.e.	295 mm w.e.
Park Pass SL	1981 – 2017	118 m	0 m	226 m	64 m
Salisbury SL	1981 – 2017	113 m	4 m	340 m	64 m
South Cameron SL	1981 – 2017	87 m	0 m	251 m	65 m
Thurneyson SL	1981 – 2017	230 m	4 m	381 m	107 m
Vertebrae12 SL	1981 – 2017	200 m	12 m	381 m	104 m
Ridge SL	1981 – 2017	72 m	0 m	223 m	67 m
Glenmary SL	1981 – 2017	137 m	0 m	343 m	94 m
Vertebrae25 SL	1981 – 2017	40 m	0 m	163 m	46 m

We also investigate the model performance, particularly the ability of the model to accurately simulate mass balance and snowlines. For the two glaciers with measured mass balance, for which we have observational uncertainty, the majority of years have overlap between the parameter suite and the confidence interval of the observed value (Figs. S.4a,c & S.5a,c). For Brewster Glacier, 8 of 13 measurements with uncertainties (62%) fall within the parameter suite (Figs. S.4a & S.5a), and for Rolleston glacier, 5 of 6 measurements with uncertainties (83%) fall within the parameter suite (Figs. S.4c & S.5). However, fewer snowline measurements fall within the parameter suite for almost all glaciers. For each of the ten glaciers, between 35% and 65% of measurements fall within the parameter suites, with an average of ∼50% for all ten glaciers.

Finally, we explore the conservation of variance in our modeled data compared with measurements. We find that for most glaciers, the modeled variance fits the measured variance. However, one glacier (Brewster Glacier) has a lower modeled variance than measured variance, for both mass balance and snowlines. This is interesting as Brewster Glacier is the only glacier with climate station data, so the precipitation and temperature are not varied as parameters (for all other glaciers, T_{adj} and P_{adj} parameters are varied in the calibrations). Previous work has shown that increasing precipitation in mass balance models results in larger model variance[17]. We test the influence of increasing variance in our model output. Brewster Glacier measured mass balance variance is 15% greater than modeled mass balance variance. We therefore create new histograms of natural and present climate mass balances with the same mean, but a standard deviation of 15% greater than the original standard deviation. With the increased variance, increases in likelihood of extreme mass loss occurring are 17 times in 2011, and 32 times in 2018. For both years, attributions calculated with these new, higher variance histograms are within the uncertainties of the attributions calculated with the original variance (Fig. 3). Brewster Glacier modeled snowlines also have a lower variance than the measured snowlines. However, again, we show that increasing the variance in the attribution histograms, increases in likelihood of extreme mass loss occurring are 12 times in 2011, and 13 times in 2018. For both years, attributions calculated with these new, higher variance histograms are within the uncertainties of the attributions calculated with the original variance.

Shortwave radiation. We also test whether including shortwave radiation in our melt calculation improves simulated mass balances and snowlines. Calculating mass balance with a traditional temperature index model for Brewster Glacier, the lowest RMSE between measured and modeled mass balance is 600 mm w.e., while the lowest RMSE with radiation added is 561 mm w.e, resulting in a 7% reduction in RMSE with shortwave radiation included in the melt calculation. For snowline elevations, the lowest RMSE between measured and modeled snowline elevations is 107 m, while the lowest RMSE with radiation is 79 m, resulting in a 26% reduction in RMSE with radiation included in the melt calculation. This improvement in both mass balance and snowline RMSEs shows the value in including radiation in our melt calculation.

Glacier geometry. All simulations are performed using glacier surface elevation from 2018 DEMs and glacier outlines from 2018 orthophoto mosaics. In a climate without anthropogenic forcing, glaciers would likely not be as small as they are in the present day. However, we are not looking to simulate glacier mass balance for different time periods: the HistoricalNat and control runs represent climate without anthropogenic forcing, not past climate. Instead we are looking to answer whether the extreme mass loss measured in 2011 and 2018 could have occurred in a climate without anthropogenic forcing. Therefore, it makes sense to use the present-day glacier geometries.

Because we are calculating attributions for glaciers in 2011 and 2018, we test whether calculated attributions for 2011 are biased by using the 2018 glacier geometries. We test this for Brewster Glacier mass loss in 2011, using both the 2011 and 2018 glacier geometries. The model was calibrated for both geometries separately, with a suite of nine parameter combinations selected for each. We find that the anthropogenic influence on Brewster Glacier mass loss in 2011 does not change between the 2011 and 2018 geometries. Using both geometries, there is a 0% (0 – 0.1%) chance of the measured 2011 mass balance occurring in a natural climate. In a climate with anthropogenic influence, there is a 3.3% (2.2 – 5.0%) chance of occurring using the 2011 geometry, and 3.2% (2.1 – 5.0%) chance using the 2018 geometry. Values are reported as the mean of the CMIP5 and CESM ensembles and the suite of varying parameters, with the range among both the ensembles and simulations varying parameters. Note that these values include uncertainty from a subset of parameter suites, but not the inherent model error, as the goal was only to compare 2011 and 2018 geometries.

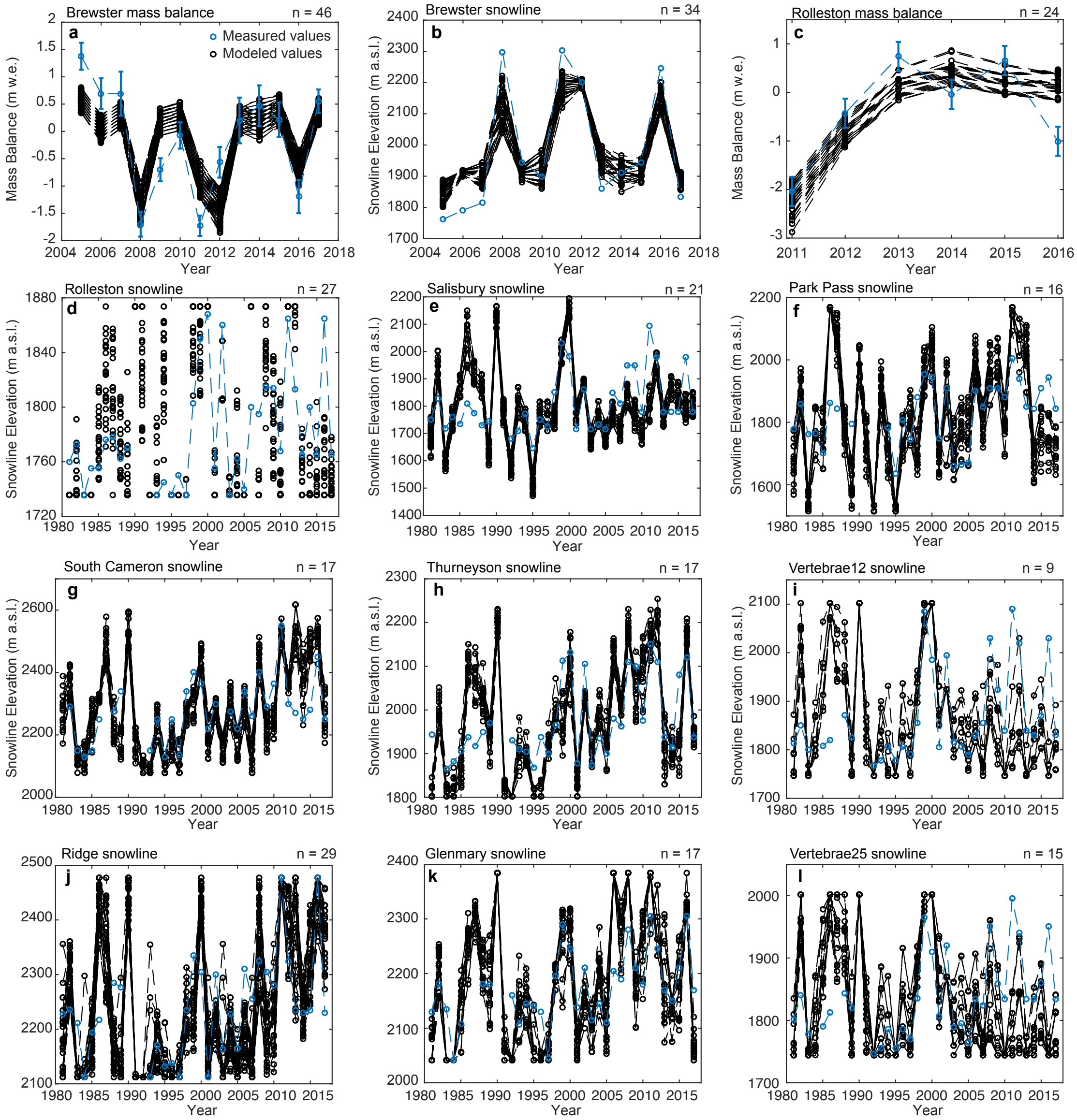

Figure S.4. Calibration results for each glacier. Measured data (blue) and modeled data (black) resulting in best fits for each glacier (n refers to the number of parameter combinations shown and used in the suite of parameter combinations for the attribution calculation).

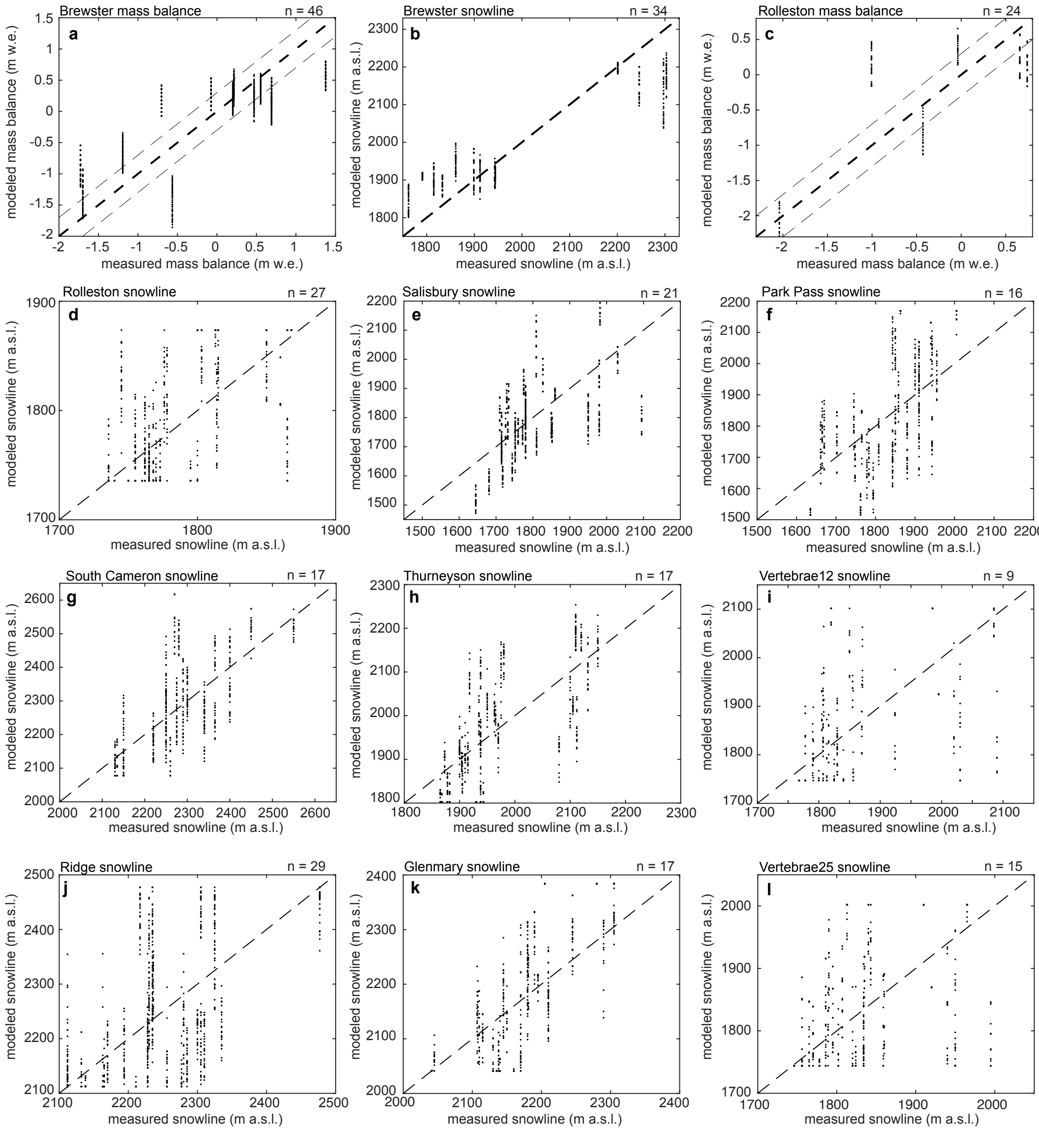

Figure S.5. Calibration results for each glacier. Model calibration scatter plots showing measured versus modeled mass balance for a) Brewster and c) Rolleston Glaciers, and snowlines for all glaciers. Each scatter plot includes all parameter combinations used in the suite of parameters for the attribution calculation (indicated by n).

	Measurement	Natural CMIP5	Natural CESM	Current CMIP5	Current CESM	Likelihood CMIP5	Likelihood CESM
Brewster mass balance	-1728 mm w.e. ('11)	0.13%	0.04%	4.6%	2.7%	35	71
	-2217 mm w.e. ('18)	<0.01%	<0.01%	1.2%	0.32%	490	290
Rolleston mass balance	-2039 mm w.e. ('11)	0.50%	0.21%	8.4%	5.5%	17	27
	-2456 mm w.e. ('18)	0.15%	0.04%	3.7%	2.0%	25	50
Brewster ELA	2303 m ('11)	0.35%	0.12%	2.5%	1.4%	7	12
	2311 m ('18)	0.30%	0.10%	2.2%	1.1%	7	12
Rolleston ELA	1857 m ('11)	7.2%	5.4%	41%	37%	6	7
	1847 m ('18)	9.2%	7.2%	49%	45%	5	6
Salisbury ELA	not measured	N/A	N/A	N/A	N/A	N/A	N/A
	1967 m ('18)	3.6%	1.6%	31%	24%	9	16
Park Pass ELA	2040 m ('11)	4.0%	2.6%	25%	26%	6	10
	1966 m ('18)	9.6%	7.1%	40%	43%	4	6
S. Cameron ELA	2395 m ('11)	11%	9.4%	45%	42%	4	5
	2426 m ('18)	8.2%	6.1%	38%	34%	5	6
Thurneyson ELA	2125 m ('11)	8.2%	6.2%	30%	30%	4	5
	2232 m ('18)	1.1%	0.4%	7.3%	6.0%	7	14
Vertebrae 12 ELA	2094 m ('11)	1.3%	1.0%	15%	11%	11	11
	2086 m ('18)	1.7%	1.2%	19%	13%	11	11
Ridge ELA	2436 m ('11)	3.3%	1.8%	18%	13%	5	7
	2434 m ('18)	3.4%	1.9%	18%	13%	5	7
Glenmary ELA	2297 m ('11)	14%	13%	36%	37%	3	3
	2327 m ('18)	9.4%	7.7%	27%	27%	3	3
Vertebrae 25 ELA	>2002 m ('11)	1.9%	1.5%	18%	14%	9	9
	>2002 m ('18)	1.9%	1.5%	18%	14%	9	9

Table S.4. **Probabilities and likelihoods of high mass loss shown by model ensemble.** The percentage of years with equal or lower mass balance, and equal or higher snowlines, for individual glaciers in the natural world and the present world, as well as the change in likelihood (ratio of present percentages to natural percentages). The mean value from the suite of model parameters is shown. Values of <0.01% represent values that would be 0 when rounded, but are still greater than 0%. This is to differentiate between probabilities that are absolutely 0% and probabilities that are greater than 0%.

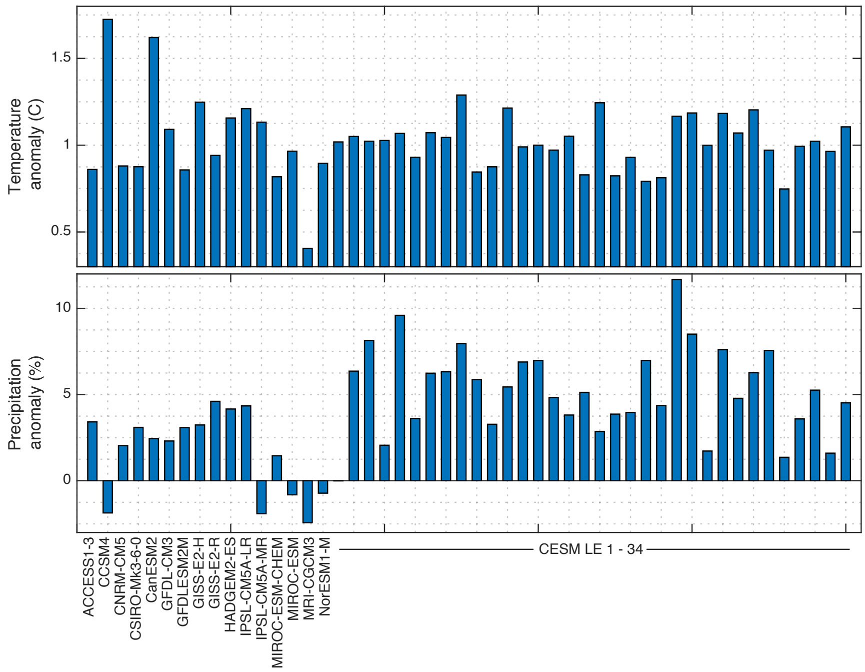

Figure S.6. Changes in climate in GCM output. Temperature (top; °C) and precipitation (bottom; %) changes for present-world scenarios compared with natural-world scenarios. For each model or CESM ensemble member, values are averaged for all ten glaciers. Present world is defined as RCP8.5 (April 2006 – March 2026) for both CMIP5 and CESM. The natural world is defined as HistoricalNat (April 1901 – March 2005) for CMIP5 and the CESM LE control run (April year 1 – March year 1800) for CESM.